ONLINE DISINFORMATION MONITORING SYSTEM

01
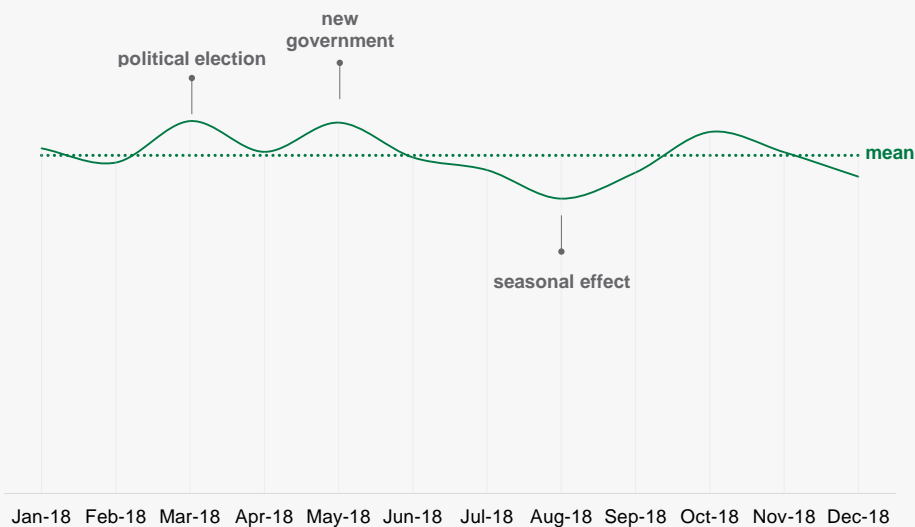19

DEPARTMENT OF ECONOMICS AND STATISTICS

AGCOM

Index

Information and disinformation production

NEWS PRODUCTION IN ITALY

Monthly trend of the news production in Italy (2018)

political election

new government

seasonal effect

mean

Jan-18 Feb-18 Mar-18 Apr-18 May-18 Jun-18 Jul-18 Aug-18 Sep-18 Oct-18 Nov-18 Dec-18

It is possible to track the effects of at least two factors on the quantity of the "news supply", produced by the media:
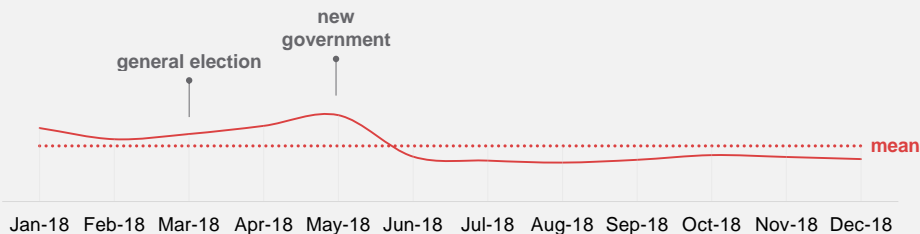
- seasonality, that occurs in the reduction suffered by the information production in summer

- political cycle, that occurs in the increase of news production during the electoral campaigns, in the last general election in Italy (March 2018) and in the subsequent formation of the new government

ONLINE DISINFORMATION PRODUCED IN ITALY

The national system suffers the presence of an amount of disinformation which:

- seems to have stabilized on an averagely higher value compared to previous years

- reached the maximum level during the last political cycle (general election and the formation of the new government)

Monthly trend of the disinformation production in Italy (2018)

general election

new government

mean

Jan-18 Feb-18 Mar-18 Apr-18 May-18 Jun-18 Jul-18 Aug-18 Sep-18 Oct-18 Nov-18 Dec-18
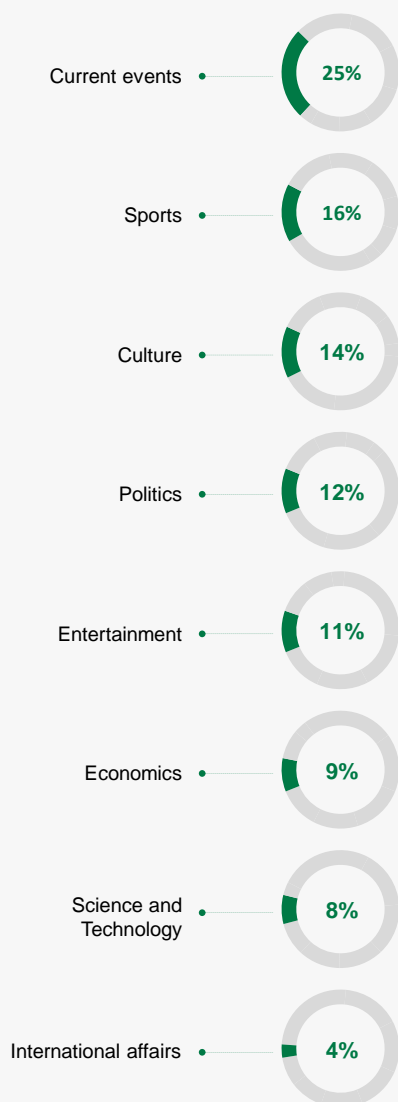
68% in the average month 2018

ONLINE CONTENTS
IN THE INFORMATION SYSTEM

8% in the average month 2018

DISINFORMATION
ON ONLINE CONTENTS

Information and disinformation categories

Supply distribution of information, by category (2018)

Current events — **25%**

Sports — **16%**

Culture — **14%**

Politics — **12%**

Entertainment — **11%**

Economics — **9%**

Science and Technology — **8%**

International affairs — **4%**

Supply distribution of disinformation, by category (2018)

34% — Current events

19% — Politics

18% — Science and Technology

10% — Entertainment

7% — Culture

6% — International affairs
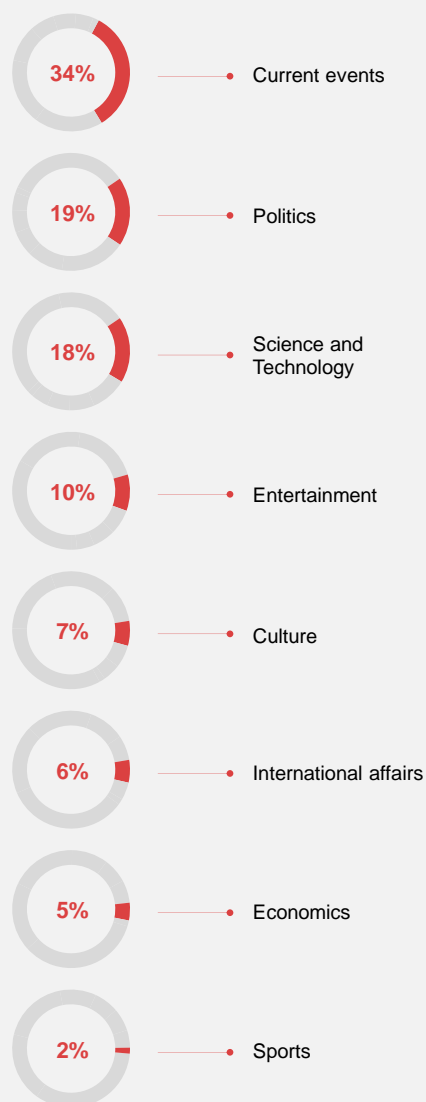
5% — Economics

2% — Sports

INFORMATION CATEGORIES

The analysis of the information produced in 2018 in Italy shows:

- the prevalence of contents concerning current events (which represent a quarter of the total) and a coverage of the hard news (current events, politics and international affairs) that overall reaches 41%

- the large space also attributed to sports news (16%), culture (14%) and entertainment (11%)

ONLINE DISINFORMATION CATEGORIES

The distribution by category of online disinformation produced in 2018 in Italy shows:

- the preponderance of contents concerning the most relevant subjects from the point of view of pluralism, such as current events and politics, jointly equal to 53%

- the low presence of sports news (2%), more easily verifiable by citizens and, therefore, less suitable for being falsified

LACK OF SCIENTIFIC INFORMATION

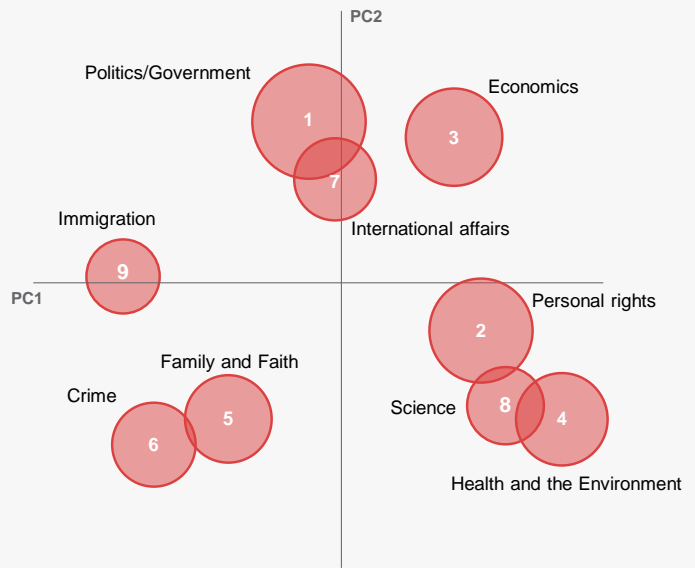## ABUNDANCE OF SCIENTIFIC DISINFORMATION

Topics of disinformation (before and after the 2018 elections)

MAIN TOPICS OF ONLINE DISINFORMATION BEFORE AND AFTER THE GENERAL ELECTION

From the textual analysis of the content of the disinformation websites (topic modeling), which allows to obtain groups of terms frequently co-occurring within the documents and identifying macro-themes, it emerges that:

- the most popular specific topics during the election period in Italy concern news on politics and government affairs, personal rights, economic issues, health and the environment, family and faith, crime, international affairs, science, immigration

Main topics of online disinformation (January-August 2018)

The list of the 30 most frequent terms referable to each topic is shown in Appendix I

Top-30 most salient terms in the disinformation websites (January-August 2018)
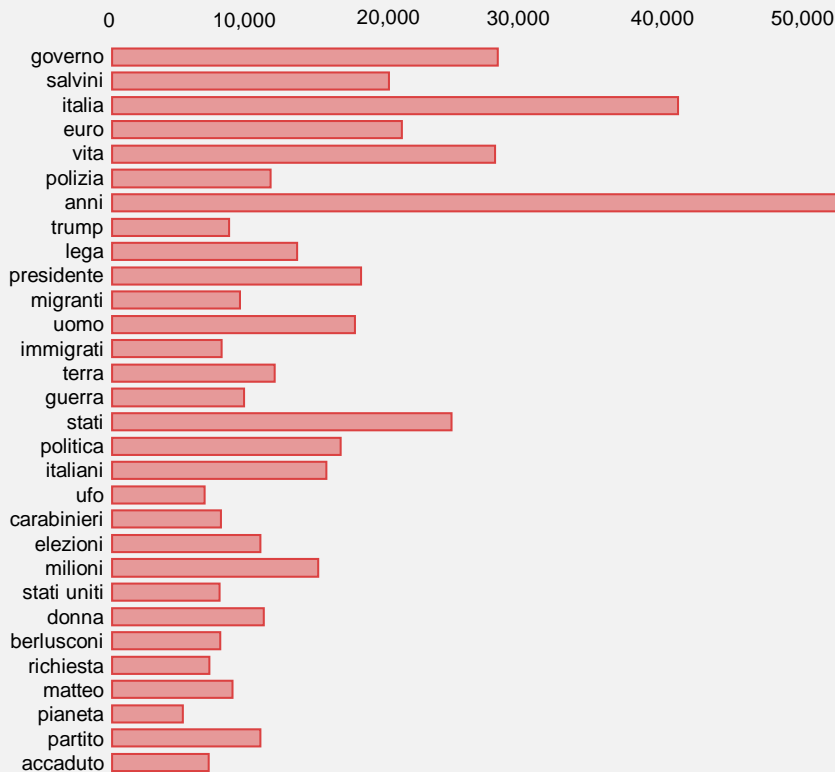
SALIENT TERMS OF THE ONLINE DISINFORMATION BEFORE AND AFTER THE GENERAL ELECTION

The analysis of the frequency distribution of all the terms contained in the disinformation websites allows to observe how:

- many of the terms that appear among the 30 most frequent in the first 8 months of 2018 are related to hard news on: political exponents and institutions, facts of crime and law enforcement, immigration, international events

- the other salient terms of disinformation in the reference period appear to be connected above all to pseudoscientific theories of planets and alien presences

DISINFORMATION STRATEGIES ARE BASED MOSTLY ON **DIVISIVE ISSUES** WITH A **STRONG EMOTIONAL IMPACT**

Topics of disinformation (second period 2018)

Main topics of online disinformation (October - December 2018)

PC2

Society

Crime

5

3

Science and Health

Politics/Government

4

1

PC1

Economics

2
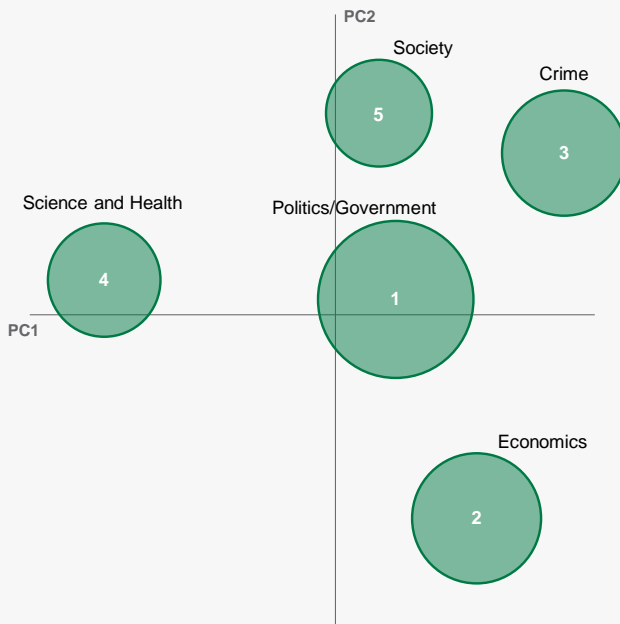
MAIN TOPICS OF ONLINE DISINFORMATION IN THE SECOND PERIOD 2018

The analysis, realized through a topic model, of the textual content produced by the disinformation websites in the last 3 months of 2018 in Italy shows:

- the emergence of 5 main themes, which, similarly to the first part of the year, concern the political, government and economic events in Italy; the crime news; science and health; and society
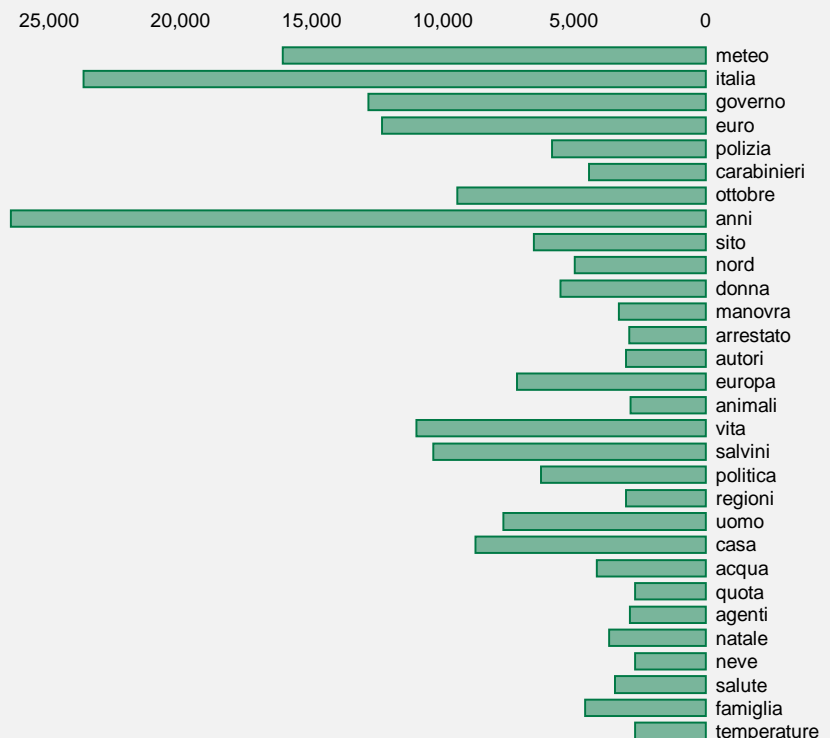
The list of the 30 most frequent terms referable to each topic is shown in Appendix II

SALIENT TERMS OF THE ONLINE DISINFORMATION IN THE SECOND PERIOD 2018

The 30 most recurring terms in the entire content of the disinformation websites, in the last quarter of 2018, appear to be linked in particular to:

- news concerning government policies and institutions, crime and police intervention, economic measures

- health, environmental and climatic conditions

Top-30 most salient terms in the disinformation websites (October - December 2018)

25,000	20,000	15,000	10,000	5,000	0

meteo
italia
governo
euro
polizia
carabinieri
ottobre
anni
sito
nord
donna
manovra
arrestato
autori
europa
animali
vita
salvini
politica
regioni
uomo
casa
acqua
quota
agenti
natale
neve
salute
famiglia
temperature

THE DISINFORMATION WEBSITES MAINLY DISSEMINATE NEWS ON **POLARIZING ISSUES**, SUITABLE TO BECOME THE OBJECT OF **VIRAL PROPAGATION** THROUGH SOCIAL NETWORKS (AND OTHER ONLINE PLATFORMS)
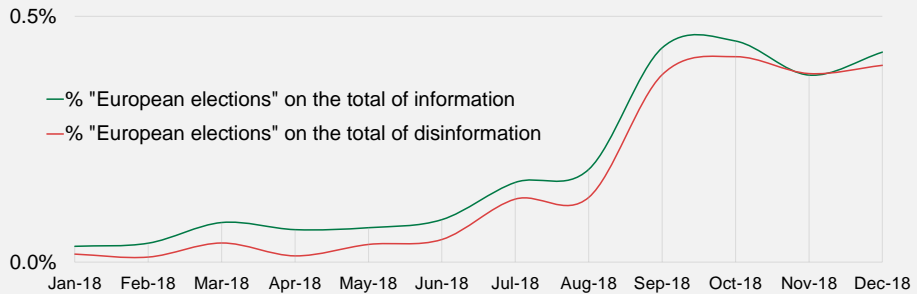
Information and disinformation on the European elections

NEWS-MEDIA COVERAGE OF THE EUROPEAN ELECTIONS IN ITALY

Comparing the information and disinformation production concerning the European elections, it emerges that:

- the incidence of such content on its respective overall supply of information and disinformation assumes similar monthly values and tendencies

- the space dedicated to the European elections both from information sources and disinformation shows an evident growth in the second half of 2018, settling on significantly higher values, especially since September

Monthly incidence of the 'European elections' topic on the total information and disinformation (2018)

— % "European elections" on the total of information
— % "European elections" on the total of disinformation

Monthly incidence of the 'European elections' topic on the total content produced by single media (2018)

Television
Disinformation websites
Newspapers

SPACE DEDICATED TO THE EUROPEAN ELECTIONS IN ITALY BY SINGLE MEDIA
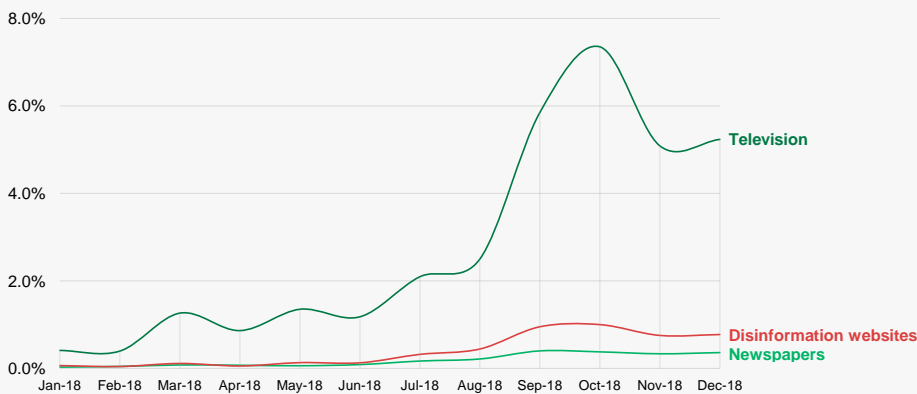
By focusing the analysis on single media, it is found that, in 2018:

- the (national) television exhibits the largest portions of space dedicated to information concerning the European elections within its own programming

- out of the total supply produced by the media, the disinformation websites, on average, devote more content to the European elections than newspapers (and any other online source)

7% ONLINE CONTENT REGARDING THE EUROPEAN ELECTIONS PRODUCED BY DISINFORMATION SOURCES (DISPLAYS AND SOCIAL NETWORK PAGES / ACCOUNTS), IN THE AVERAGE MONTH

Importance of the topics of European relevance among the Italians

Most important topics of European relevance among the Italian citizens
(multiple choice, maximum 3 answers)

Topic	% population
Immigration	59%
Economic situation	49%
Unemployment	47%
Climate change	33%
Terrorism	24%
Crime	16%
Taxation	12%
Pensions	11%
Public finances	10%
Inflation	4%

■ % population
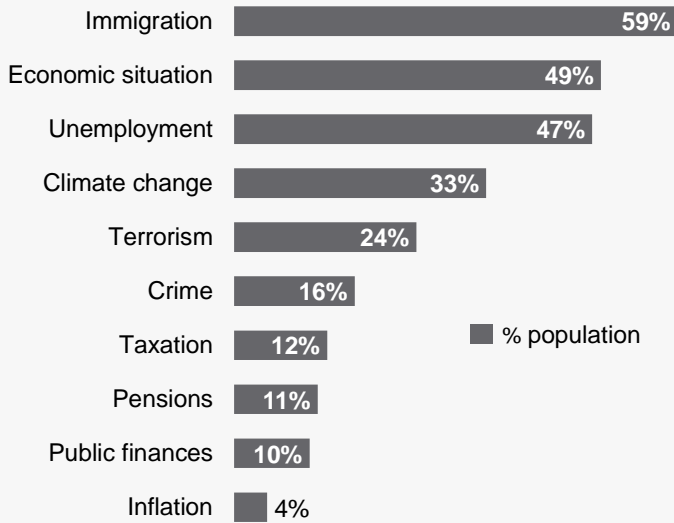
Source: Agcom elaborations of SWG data, January 2019

IMPORTANCE GIVEN BY ITALIANS TO TOPICS OF EUROPEAN RELEVANCE

Based on a SWG investigation commissioned by Agcom on a sample of 1,358 individual aged from 14 to 74 years old, we found that:
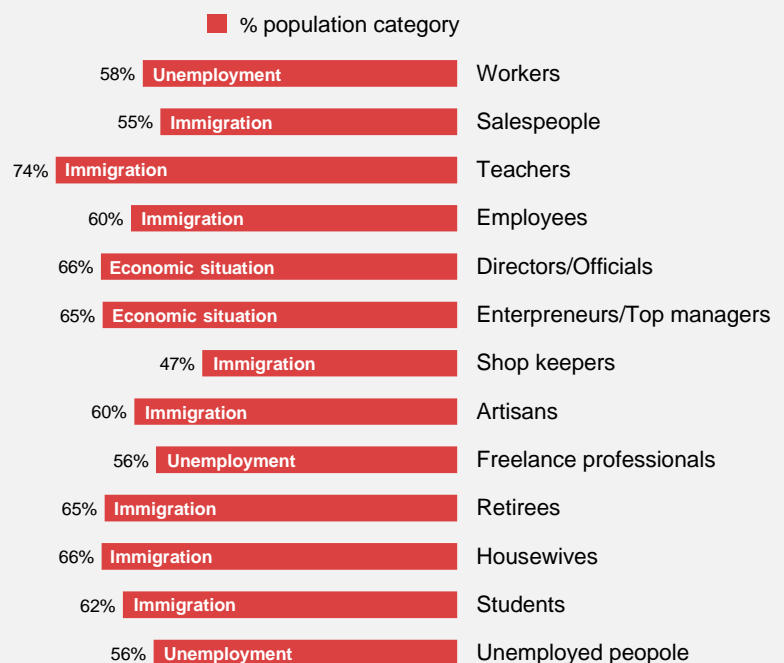
- as for the topics of EU relevance identified by Eurobarometer, immigration is in the top three list among the 60% of citizens, followed by national economic situation and unemployement that account for 50%

- also climate change and terrorism are among the top five topics considered as having most relevance for the EU, with a score of 33% e and 24% respectively

MOST CONCERNING TOPIC OF EUROPEAN RELEVANCE AMONG SINGLE CITIZENS' GROUP

Based on the importance given to topics of European relevance by each professional category of citizens, we have observed that:

- immigration is the most frequently indicated topic by salespeople, teachers, employees, shop keepers, artisans, students, housewives, and retirees

- the economic situation in Italy is the most relevant topic among the most remunerated professional categories of citizens, that are directors, officials, top managers and enterpreneurs

- unemployment is the most concerning topic not only among unemployed peopole but also among workers and freelance professionals.

Top concerning topic, by professional category
(multiple choice, maximum 3 answers)

■ % population category

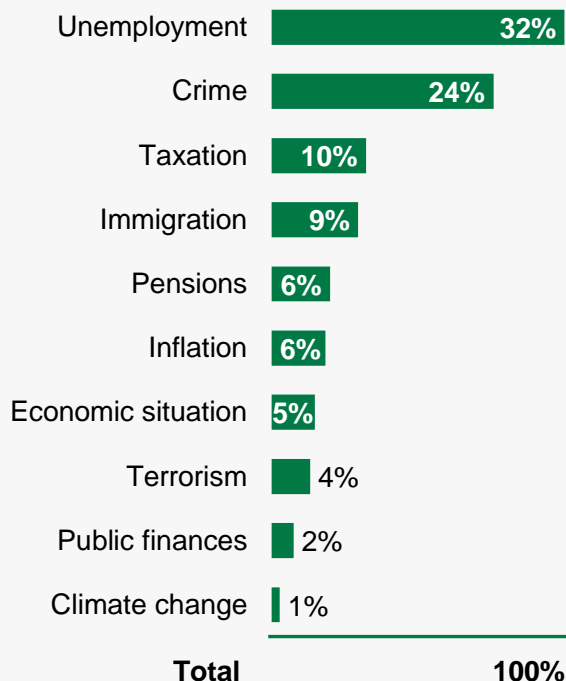| %	Top topic	Category
58%	Unemployment	Workers
55%	Immigration	Salespeople
74%	Immigration	Teachers
60%	Immigration	Employees
66%	Economic situation	Directors/Officials
65%	Economic situation	Enterpreneurs/Top managers
47%	Immigration	Shop keepers
60%	Immigration	Artisans
56%	Unemployment	Freelance professionals
65%	Immigration	Retirees
66%	Immigration	Housewives
62%	Immigration	Students
56%	Unemployment	Unemployed peopole

Source: Agcom elaborations of SWG data, January 2019

IMMIGRATION IS THE TOPIC OF EUROPEAN RELEVANCE THAT ITALIANS ARE MOST CONCERNED ABOUT

ECONOMIC SITUATION AND UNEMPLOYMENT ARE THE OTHER MOST CONCERNING TOPICS

Information and disinformation on topics of European relevance

Information supply distribution, by topic (2018)

Topic	%
Unemployment	32%
Crime	24%
Taxation	10%
Immigration	9%
Pensions	6%
Inflation	6%
Economic situation	5%
Terrorism	4%
Public finances	2%
Climate change	1%
Total	**100%**

Disinformation supply distribution, by topic (2018)

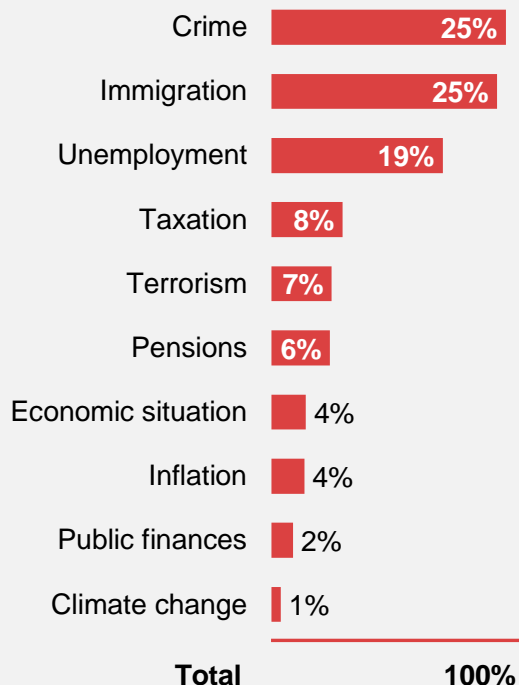| Topic	%
Crime	25%
Immigration	25%
Unemployment	19%
Taxation	8%
Terrorism	7%
Pensions	6%
Economic situation	4%
Inflation	4%
Public finances	2%
Climate change	1%
Total	**100%**

INFORMATION ON TOPICS OF EUROPEAN RELEVANCE

By focusing the analysis only on the 2018 news supply concerning te topics of European relevance, it apperars that in Italy:

- unemployment and crime are the most covered topics

- less covarage is given to news on terrorism, public finance and, most of all, climate change

ONLINE DISINFORMATION ON TOPICS OF EUROPEAN RELEVANCE

From online disinformation distribution, regarding topics of European relevance, it appears that in Italy:

- crime, immigration, unemployment are largely the most covered topics

- immigration is a quarter of the online disinformation content on topics of European relevance, whereas the corresponding percentage of information content is 9%

AMONG TOPICS OF EUROPEAN RELEVANCE THE MOST COVERED ONLINE ARE **IMMIGRATION** AND **CRIME**

60%

ONLINE CONTENT ON THE TOTAL OF CRIME

57%

ONLINE CONTENT ON THE TOTAL OF IMMIGRATION

IMMIGRATION AND **TERRORISM** ARE AMONG THE TOPICS WITH THE MOST DISINFORMATION ON THE TOTAL CONTENT ONLINE

15%

DISINFORMATION ON TOTAL ONLINE CONTENTS OF IMMIGRATION

11%

DISINFORMATION ON TOTAL ONLINE CONTENTS OF TERRORISM

Appendix I

SALIENT TERMS OF EACH TOPIC OF ONLINE DISINFORMATION BEFORE AND AFTER THE 2018 ELECTIONS

The composition of each topic of disinformation allows to obtain, based on the salient terms associated with them, indications about the specific object of the disseminated news.

In this sense, for each of the 9 main topics of disinformation, we list the 30 most frequent terms that compose them:

1. POLITICS/GOVERNMENT

- governo
- italia
- salvini
- politica
- lega
- presidente
- fatto
- partito
- italiani
- matteo
- paese
- oggi
- movimento
- berlusconi
- ministro
- elettorale
- forza
- leader
- renzi
- legge
- stelle
- anni
- voto
- elezioni
- sinistra
- politico
- politiche
- cittadini
- media
- invece

2. PERSONAL RIGHTS

- vita
- persone
- modo
- altri
- mondo
- tempo
- persona
- bambini
- altro
- avere
- spesso
- realtà
- amore
- momento
- bisogno
- viene
- mente
- volte
- possiamo
- fatto
- lavoro
- propria
- bambino
- senso
- meglio
- dire
- leggi
- importante
- corpo
- tutte

3. ECONOMICS

- euro
- milioni
- anno
- miliardi
- anni
- dati
- rispetto
- sistema
- lavoro
- società
- mercato
- grazie
- mondo
- oltre
- denaro
- apre
- aumento
- economia
- nuovi
- paesi
- condividere
- italia
- aziende
- crescita
- finestra
- livello
- modo
- viene
- numero
- banca

4. HEALTH AND THE ENVIRONMENT

- acqua
- meteo
- italia
- salute
- farmaci
- nord
- tempo
- aria
- effetti
- giorno
- giorni
- sale
- pressione
- neve
- temperature
- regioni
- pelle
- olio
- curcuma
- dieta
- qualche
- caso
- naturale
- rischio
- proprietà
- causa
- corpo
- centro
- malattie
- mese

5. FAMILY AND FAITH

- anni
- casa
- fatto
- storia
- famiglia
- padre
- quel
- giorno
- foto
- roma
- uomo
- chiesa
- donna
- qualche
- volta
- francesco
- oggi
- insieme
- nome
- san
- giorni
- morte
- gennaio
- figlio
- dice
- vita
- durante
- tanto
- febbraio
- madre

Appendix I

6. CRIME

- anni
- polizia
- uomo
- carabinieri
- donna
- ospedale
- persone
- auto
- agenti
- arrestato
- stati
- casa
- procura
- leggi
- giovane
- carcere
- tribunale
- mesi
- fatto
- state
- indagini
- violenza
- vittima
- milano
- giorni
- posto
- accusa
- scorso
- altri
- momento

7. INTERNATIONAL AFFAIRS

- trump
- guerra
- stati uniti
- usa
- siria
- presidente
- russia
- europa
- soros
- stati
- paese
- paesi
- mondo
- governo
- attacco
- militare
- francia
- armi
- america
- dichiarato
- francese
- persone
- forze
- sicurezza
- donald
- mondiale
- fatto
- media
- iran
- internazionale

8. SCIENCE

- terra
- ufo
- anni
- pianeta
- mondo
- luce
- glenn
- luna
- segnidalcielo
- sistema
- alieni
- scienziati
- giganti
- ricercatori
- altri
- stati
- ricerca
- oggetti
- esseri
- vita
- storia
- nasa
- immagini
- vicino
- solare
- oggetto
- tecnologia
- libro
- antichi
- civiltà

9. IMMIGRATION

- stati
- italia
- migranti
- immigrati
- richiesta
- accaduto
- media
- clandestini
- settimana
- profughi
- annunciato
- italiani
- accoglienza
- elezioni
- città
- account
- centro
- roma
- sindaco
- indipendenti
- satira
- asilo
- islamizzazione
- scomparse
- scl
- disattivati
- euro
- casa
- lascia
- cittadini

Appendix II

SALIENT TERMS OF EACH TOPIC OF ONLINE DISINFORMATION IN THE SECOND PERIOD 2018

For each of the 5 main topics of disinformation in the last quarter of 2018, the list of the 30 most frequent terms associated with them is shown.

In this regard, it should be noted that sometimes the composition of a thematic group may also include a significant subset of terms referring to other specific topics. For example, it is noted that within the second group of terms, mainly inherent to the economic sphere, others are related to the climatic conditions

1. POLITICS/GOVERNMENT

- fatto
- salvini
- italia
- governo
- presidente
- oggi
- politica
- mondo
- persone
- ministro
- vita
- italiani
- paese
- dire
- tempo
- partito
- anni
- matteo
- lega
- europa
- fonte
- movimento
- guerra
- dice
- forza
- volta
- sinistra
- parole
- deve
- caso

2. ECONOMICS

- meteo
- italia
- euro
- governo
- anni
- nord
- legge
- anno
- europa
- paese
- giorni
- manovra
- milioni
- oggi
- lavoro
- regioni
- tempo
- dicembre
- miliardi
- aria
- sud
- quota
- neve
- commissione
- inizio
- europea
- temperature
- bilancio
- presidente
- maltempo

3. CRIME

- anni
- ottobre
- polizia
- stati
- carabinieri
- uomo
- casa
- donna
- persone
- roma
- arrestato
- agenti
- auto
- immigrati
- euro
- ordine
- interno
- centro
- città
- fatto
- giorni
- zona
- napoli
- san
- giovane
- procura
- forze
- posto
- droga
- corso

4. SCIENCE AND HEALTH

- sito
- acqua
- autori
- salute
- animali
- tempo
- viene
- anni
- caso
- terra
- corpo
- società
- dati
- sistema
- persone
- studio
- vita
- prodotto
- ricerca
- energia
- pianeta
- articolo
- rispetto
- prodotti
- sport
- responsabili
- potrebbero
- incredibili
- rischio
- mondo

5. SOCIETY

- anni
- vita
- fatto
- casa
- famiglia
- natale
- bambini
- storia
- anno
- foto
- donna
- mondo
- donne
- tempo
- oggi
- giorno
- social
- figlio
- insieme
- genitori
- durante
- scuola
- volta
- uomo
- video
- giorni
- madre
- figli
- napoli
- padre

Methodological note

THE COMPOSITION OF THE DATABASE

The information reported in this document, unless otherwise specified, are the result of Agcom elaborations carried out on a database built from data extrapolated through the platform developed by *Volocom Technology*.

In particular, the analyzes were conducted on the entire textual content extrapolated from about 15 million documents created in Italy in 2018 by 1,800 information sources (national television and radio channels, newspapers, news agencies, websites of traditional publishers, native online news outlets, and related pages and accounts of social networks), and sources of disinformation (websites and social pages/accounts) identified as such by external subjects specialized in debunking activities. The amount of online disinformation produced in Italy was therefore estimated using a subjective methodology, i.e. considering the total number of documents created monthly by the aforementioned sources of disinformation.

The database is composed by the entire textual content of all documents produced during a day by every information and disinformation source. With document, we mean the entire article, in the case of newspapers and information websites; the transcription of a transmission segment, in the case of Tv and radio: all tweets/posts in the case of online platforms.

TOPIC MODELING

For the study of the main topics of online disinformation, analyzes were carried out on the entire content disseminated by the disinformation websites in 2018, adopting a methodological approach known as *topic modeling*.

A topic model is a statistic model for the automatic individuation of topics appearing in a collection of documents. In particular, the classification of the text together to the group of fake contents in determined topics was obtained by means of the use of a LDA-model (Latent Dirichlet Allocation) - an algorithm of not-supervised automatic learning considering the frequency and the co-occurrence of the terms used in the collection of documents. The same is at the basis of LDAvis (Sievert e Shirley, 2014), the interactive visualization system that supplies an overall vision on the identified topics (and of the way in which they differ), allowing at the same time a deep analysis of the most salient terms, associated to every one of them. In detail, every topic is represented by a circle, which amplitude shows the occurrence frequency in the entire group of the examined documents. Moreover, the topics have a position in the Cartesian coordinate system, being around the horizontal and vertical axes, representing the two main dimensions that explain the variability of the data.

FURTHER INFORMATION

For more information on the methodological aspects, see Agcom Report (2018), *News vs. fake in the information system*

For more technical information and definitions regarding online disinformation, see Agcom Report (2018), *Online disinformation strategies and the fake content supply chain*

For further details and updates on the Agcom's activities in the field of online disinformation, it is possible to consult the area dedicated to Technical Roundtable for safeguarding News-Media Pluralism and Fairness in the online Platforms

DEPARTMENT OF ECONOMICS AND STATISTICS