ONLINE DISINFORMATION MONITORING SYSTEM

AGCOM

03
19

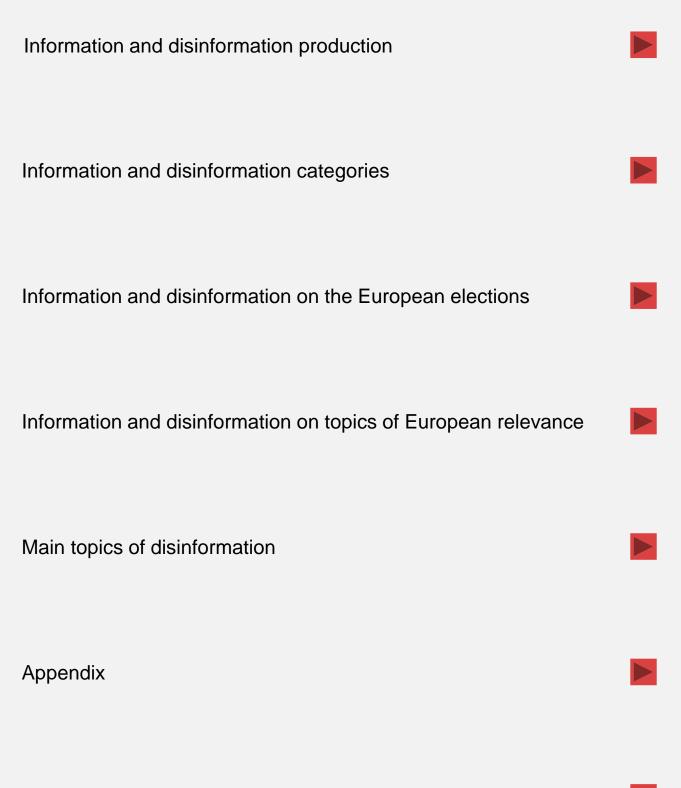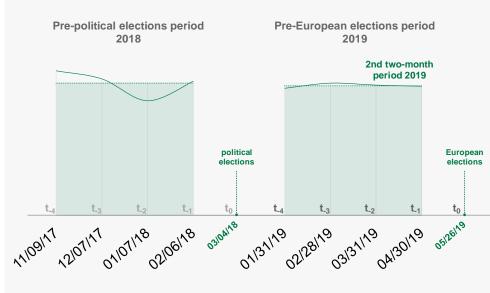**DEPARTMENT OF ECONOMICS AND STATISTICS**

Index

Information and disinformation production

Total news production on an average day

Pre-political elections period 2018

Pre-European elections period 2019

2nd two-month period 2019

political elections

European elections

t$_{-4}$ t$_{-3}$ t$_{-2}$ t$_{-1}$ t$_{0}$ t$_{-4}$ t$_{-3}$ t$_{-2}$ t$_{-1}$ t$_{0}$

11/09/17   12/07/17   01/07/18   02/06/18   03/04/18   01/31/19   02/28/19   03/31/19   04/30/19   05/26/19

The analysis of the overall news production in Italy shows:

- an almost constant trend in the second two-month period 2019 compared to the previous two-month period

- a news production on the average day of the pre-European elections period 2019 slightly lower than the news production in the pre-political elections period 2018
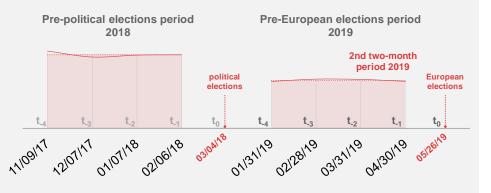
See information on the European elections

ONLINE DISINFORMATION PRODUCED IN ITALY

With regard to the overall disinformation produced in Italy (by disinformation websites, social pages and accounts), we find:

- a disinformation production that on an average day of the two-month period March-April 2019 remained at similar values than the average day of January and February

- an average daily amount of online disinformation content significantly lower in the first four months of 2019 compared to the corresponding period preceding the political elections of 2018

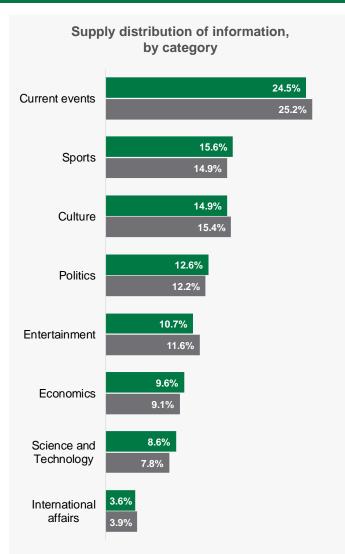### Total disinformation production on an average day

Pre-political elections period 2018

Pre-European elections period 2019

2nd two-month period 2019

political elections

European elections

t$_{-4}$ t$_{-3}$ t$_{-2}$ t$_{-1}$ t$_{0}$ t$_{-4}$ t$_{-3}$ t$_{-2}$ t$_{-1}$ t$_{0}$

11/09/17   12/07/17   01/07/18   02/06/18   03/04/18   01/31/19   02/28/19   03/31/19   04/30/19   05/26/19

See disinformation on the European elections

7% on an average day of the second two-month period 2019

DISINFORMATION ON ONLINE CONTENTS

= to the first two-month period 2019

ONLINE DISINFORMATION IN THE SECOND TWO-MONTHS PERIOD
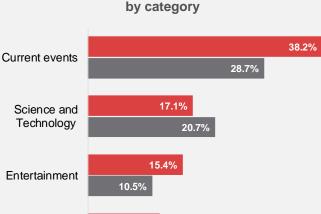
Information and disinformation categories

Supply distribution of information, by category

Category	Pre-European elections period 2019	Pre-political elections period 2018
Current events	24.5%	25.2%
Sports	15.6%	14.9%
Culture	14.9%	15.4%
Politics	12.6%	12.2%
Entertainment	10.7%	11.6%
Economics	9.6%	9.1%
Science and Technology	8.6%	7.8%
International affairs	3.6%	3.9%

■ Pre-European elections period 2019 (03/01/19 - 04/30/19)
■ Pre-political elections period 2018 (12/08/17 - 02/06/18)

Supply distribution of online disinformation, by category

Category	Pre-European elections period 2019	Pre-political elections period 2018
Current events	38.2%	28.7%
Science and Technology	17.1%	20.7%
Entertainment	15.4%	10.5%
Politics	11.8%	18.6%
Culture	7.2%	8.6%
International affairs	4.7%	6.9%
Economics	3.9%	5.0%
Sports	1.7%	0.9%

■ Pre-European elections period 2019 (03/01/19 - 04/30/19)
■ Pre-political elections period 2018 (12/08/17 - 02/06/18)
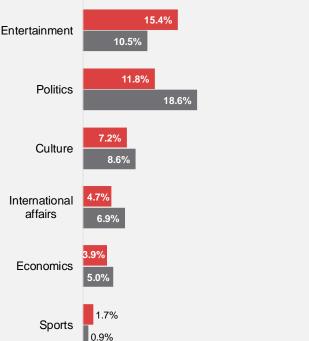
INFORMATION CATEGORIES

Examining the supply distribution of information by category, it is noted that, in the second two-month period of 2019:

- ''current events'' remains the most treated, representing the object of about a quarter of the informative contents, and the incidence of the other hard news categories (political and international affairs) is on the whole equal to 16%

- there are no significant changes compared to the same pre-election period of the previous year

ONLINE DISINFORMATION CATEGORIES

The distribution by category of online disinformation articles in the period March-April 2019 shows:

- a preponderance of contents dedicated to "current events", which, together with those of politics, constitute half of the total. The weight of scientific disinformation is also relevant (17%)

- some differences with respect to the distribution of the corresponding period before the 2018 elections, where there is a notable increase for "current events" and "entertainment" in the face of a reduction, especially of politics

PERSISTENT **LACK** OF **INFORMATION** ON **SCIENTIFIC** ARGUMENTS

ABUNDANCE OF ONLINE **DISINFORMATION** ON **CURRENT EVENTS**

Information and disinformation on the European elections

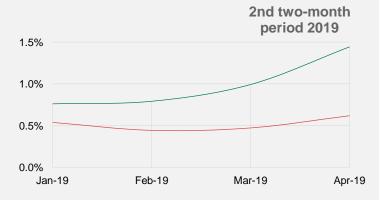## NEWS-MEDIA COVERAGE OF THE EUROPEAN ELECTIONS IN ITALY

We observe that, in the second two-month period 2019:

- the space devoted to information on the European elections, while showing a growing trend, remains below 1.5% of the total

- overall, the specific ''European elections'' topic assumes, for the disinformation sources (websites and social pages/accounts) a growing incidence, which does not reach 1% however

- disinformation contents produced on the European elections, although representing a small percentage of the total, increased by 5% in March (compared to February) and by 21% in April (compared to March)
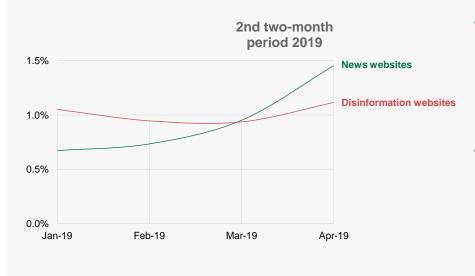
Incidence of the ''European elections'' topic on the total information and disinformation

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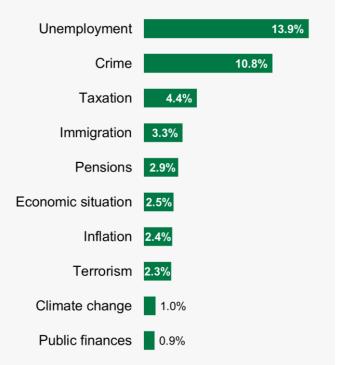

2nd two-month period 2019

SPACE DEDICATED TO THE EUROPEAN ELECTIONS IN ITALY BY WEBSITES

Focusing the analysis on the websites, it is found as:

- for disinformation websites, the incidence of articles dealing directly with the "European elections" topic exceeds 1% of the total content produced in April

- unlike the previous months, in the second two-month period of 2019 (in particular in April), the weight of the ''European elections'' topic assumes for the disinformation websites a value lower than that of the news websites

Incidence of the ''European elections'' topic on the total content produced by websites

2nd two-month period 2019

News websites
Disinformation websites

LITTLE SPACE DEDICATED TO THE **EUROPEAN ELECTIONS** BY INFORMATION AND DISINFORMATION SOURCES ON AN AVERAGE DAY OF THE SECOND TWO-MONTH PERIOD 2019

Information and disinformation on topics of European relevance

Incidence of topics of European relevance on the total information (second two-month period 2019)

Topic	Value
Unemployment	13.9%
Crime	10.8%
Taxation	4.4%
Immigration	3.3%
Pensions	2.9%
Economic situation	2.5%
Inflation	2.4%
Terrorism	2.3%
Climate change	1.0%
Public finances	0.9%

Incidence of topics of European relevance on the total disinformation (second two-month period 2019)

Topic	Value
Crime	10.4%
Unemployment	7.0%
Immigration	6.0%
Terrorism	2.4%
Taxation	2.0%
Pensions	1.5%
Inflation	1.1%
Economic situation	1.1%
Climate change	0.7%
Public finances	0.4%

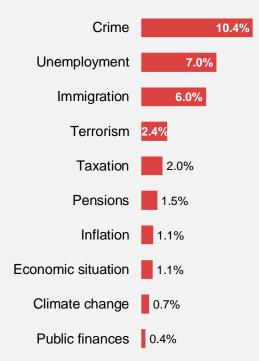## INFORMATION ON TOPICS OF EUROPEAN RELEVANCE

By focusing on the themes of European relevance identified by Eurobarometer, we find that in Italy:

- in the second two-month period of 2019, as in the previous one, unemployment and crime remain the themes with the highest incidence on the total information content (respectively 14% and 11%)

- taxation (4%), immigration (3%) and pensions (3%) follow. Less space is devoted to information on climate change (albeit growing) and to the state of public finances

ONLINE DISINFORMATION ON TOPICS OF EUROPEAN RELEVANCE

With regard to online disinformation produced on topics of European relevance in the 2019 March-April period, it is worth noting that:

- crime remains the theme most dealt with by sources of disinformation in Italy, with an incidence of 10% on the total

- the weight of immigration (6%) decreases, which falls in third place, ranking behind unemployment (7% of total)

- the incidence of content dedicated to terrorism is growing, exceeding 2% of the total

AMONG THE TOPICS OF EUROPEAN RELEVANCE:

UNEMPLOYMENT REMAINS THE MOST COVERED BY TV, RADIO, NEWSPAPERS AND NEWS WEBSITES

CRIME IS THE MOST TREATED BY THE DISINFORMATION WEBSITES AND BY THE SOCIAL SOURCES OF INFORMATION AND DISINFORMATION

IN THE SECOND TWO MONTHS OF 2019, ONLINE DISINFORMATION FOCUSES **MORE** ON **CRIME** AND **UNEMPLOYMENT**, **LESS** ON **IMMIGRATION**
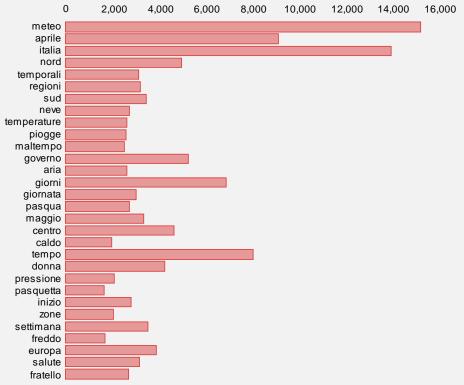
Main topics of disinformation

MAIN TOPICS OF ONLINE DISINFORMATION

Textual analysis (*topic modeling*) of the disinformation websites' content allows obtaining groups of frequently co-occurring terms in the analyzed documents, and identifying macro-themes. It emerges that:
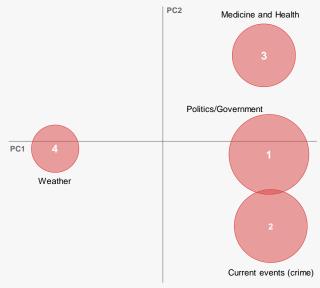
- in the second two-month period of 2019, the 4 main topics covered are related to politics and government; crime; medicine and health; weather. These topics are typically divisive and have a strong emotional impact, linked to facts and phenomena of particular interest to citizens or suitable to arouse apprehension

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

Main topics of online disinformation (March-April 2019)

Top-30 most salient terms in the disinformation websites (March-April 2019)

SALIENT TERMS OF ONLINE DISINFORMATION

The frequency distribution of the terms contained in the articles of the disinformation websites shows how:

- many of the 30 most frequent terms are related to weather forecasts (e.g. "meteo", "tempo") and adverse weather conditions (e.g. "temporali", "piogge", "neve", "maltempo")

- among the first 15 terms that appear most often in disinformation websites, there are those indicative of issues of particular relevance in terms of pluralism and personal safeguard (such as "governo", "europa", "salute")

POLITICS/GOVERNMENT, CRIME AND **MEDICINE/HEALTH** ARE THE MAIN TOPICS OF ONLINE DISINFORMATION IN THE SECON TWO-MONTH PERIOD 2019

Appendix

SALIENT TERMS OF THE MAIN TOPICS OF ONLINE DISINFORMATION

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

In this sense, for each of the 4 main topics of disinformation in the second two-month period 2019, we list the 30 most frequent terms that compose them:

1. POLITICS/GOVERNMENT	2. CURRENT EVENTS (CRIME)
• italia	• anni
• governo	• fatto
• euro	• casa
• anni	• donna
• fatto	• uomo
• salvini	• vita
• oggi	• polizia
• nazionale	• momento
• presidente	• stati
• politica	• fratello
• legge	• donne
• ministro	• durante
• roma	• puntata
• paese	• napoli
• cittadini	• uomini
• italiani	• carabinieri
• lega	• giorni
• momento	• foto
• caso	• persone
• lavoro	• giovane
• stati	• volta
• europa	• figlio
• milioni	• famiglia
• anno	• insieme
• rimanere	• tempo
• persone	• oggi
• movimento	• aprile
• guerra	• social
• partito	• giorno
• città	• ragazza

3. MEDICINE AND HEALTH	4. WEATHER
• mondo	• meteo
• anni	• aprile
• vita	• italia
• salute	• nord
• sistema	• temporali
• lavoro	• regioni
• persone	• sud
• dieta	• neve
• cura	• temperature
• corpo	• giorni
• controllo	• piogge
• vera	• maltempo
• tempo	• tempo
• causa	• aria
• blog	• giornata
• grazie	• centro
• medicina	• pasqua
• vaccini	• maggio
• oroscopo	• caldo
• malattia	• settimana
• viene	• inizio
• prevenzione	• pressione
• terra	• pasquetta
• cancro	• zone
• giorno	• freddo
• oggi	• metri
• spesso	• previsioni
• acqua	• mese
• articolo	• forti
• hiv	• sardegna

Methodological note

THE COMPOSITION OF THE 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 particular, the analyzes were conducted on the entire textual content extrapolated from more than 10 million documents created in Italy (in the first four months of 2019 and in the corresponding period before the 2018 political elections) by more than 2,000 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 the second two-month period of 2019, 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

All the numbers of the Online Disinformation Monitoring System can be consulted in a specific section of the Agcom website.

www.agcom.it

ses@agcom.it

DEPARTMENT OF ECONOMICS AND STATISTICS