Misinformation During Covid-19 Era in Greece – The Dataset 'Govgr_Covid19_Misinfo'

Georgios Lamprou - Mathematician MSc.

2025-03-13

Introduction

Nowadays, the daily amount of information that is created globally could be considered astronomical. In 2018, Bernard Marr wrote an article in Forbes magazine that showed the fact that in daily frequency 2.5 quintillion data are born (Marr 2018). About 2 weeks ago, the question: "How much information is created daily on the whole world?" was asked in the famous AI model "ChatGPT". The answer was shocking: 350 billion GB of data are born every day with this number continuing to grow exponentially. So, it is logical that all this data must be managed and analyzed for the benefit of society. However, within this huge amount of information, a significant proportion is taken up by misinformation.

According to the Council of Europe, misinformation is the situation that false information is created without having a purpose of harming a social network (Wardle and Derakhshan 2017). There are also two more categories of misinformation: disinformation and malinformation. Disinformation could be characterized as the information that is created under sneaky purposes and malinformation is the usage of not false data in order to harm one single person or a whole community. A famous example of malinformation is the usage of personal data. Despite these three categories of misinformation, people accept and transfer thousands of not true information related to different topics in a daily basis.

Misinformation and Covid-19

During the Covid-19 period, the misinformation phenomenon was dramatically getting bigger through social media like X (former Twitter), TikTok, Facebook etc. (Madraki et al. 2021). A World's Health Organization (WHO) report noticed that the misinformation sentences can be distributed into 23 categories such as "alcohol", "hands' cleaning", "cold weather/snow", "protection face masks", "viruses, bacterial, antibiotics" and others. (Organization et al. 2021). In addition, a severe percentage of misinformation developed around the 5G networks and how much possible is the virus's transmission through them. Obviously, the probability of that is zero (Nsoesie et al. 2020).

Protest in USA against compulsory vaccination, January 2022 - Source: The Washington Post

The Covid-19 vaccines discovery could be marked as an important reason for misinformation growth, too. To specify this, controversial opinions and online conflicts were noticed relative to vaccines' efficiency while antivaccine protests took place in many cities around the world (Brennen et al. 2020; Gisondi et al. 2022). These protests largely

raised the issues of compulsory vaccination against Covid-19 and free choice as well as the correctness of governmental decisions and policies (Liao 2022).

Misinformation and Covid-19 in Greece

The corresponding situation in Greece did not differ at all. Even though the management of the 1st phase of Covid-19 pandemic characterized as efficient, the 2nd phase was accompanied by the spread of inaccurate information about the virus Sars-Cov-2 and the Covid-19. On the one hand, in the 1st phase, the health scientists as well as the mass media managed the pandemic with an exemplary way and the Greek people showed admirable discipline in the Covid-19 protections measures. On the other hand, the 2nd phase of pandemic in combination with the 2nd general lockdown, boosted society in psychological tiredness and irritation against the preventive measures (Aspriadis 2021). But it is truth that the misinformation intensified from the vaccination's against Covid-19 start, too. The most frequent consequences of this were the cause of negative emotions to people such as fear, anxiety and concern about vaccines efficiency and the creation of gatherings and demonstrations against compulsory Covid-19 vaccination (Giannouli et al. 2024; Raballo, Poletti, and Preti 2022).

*Protest in Greece against Compulsory Vaccination – July 2021 – Source: AP News*

The Battle of Mass Media against Misinformation about Covid-19

To tackle the misinformation phenomenon around Covid-19, several social and mass media conducted information checking with the help of human resources or the usage of appropriate algorithms (Giannouli et al. 2024). Another important effort was the fact that many news shows, invited and hosted health specialists, doctors and scientists, which were mostly all of them members of the Greek Committee of Infectious Diseases (Anwar et al. 2020). Remarkable action was also the database's construction from WHO that is called "Public Health and Social Measures in Response to COVID-19". This database provides formal data about preventive measures and policies against Covid-19 from all over the world. For the case of Greece, interesting data are gathering. In particular, the following graph shows that to create this database from the 807 records, 475 records came from governmental sources, 209 from media articles and 123 from relatives WHO's sites.
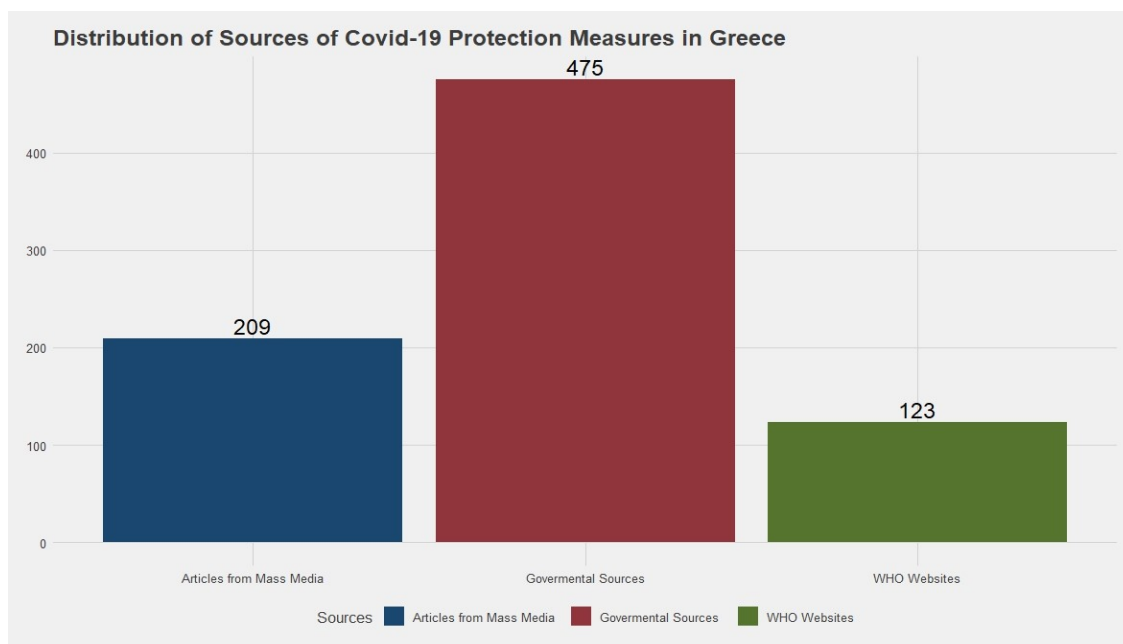
Distribution of Sources of Covid-19 Protection Measures in Greece

The above graph verifies the scientific community on the topic of reliable dataset construction. This kind of construction requires data gathering from different sources categories and not only from one (Soranno et al. 2015). Furthermore, the graph below presents the sites that were used for the creation of this WHO's database. So, the sites that used the most are those of the Greek Ministry of Health (145 records) and those of the formal governmental website for Covid-19 (131 records). In addition, 141 records came from articles by "Kathimerini" newspaper.
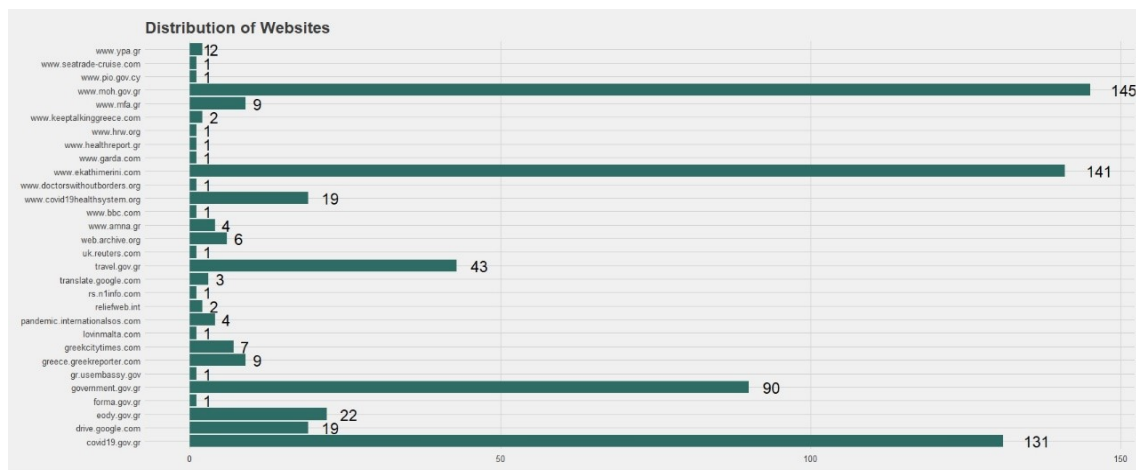
Distribution of Websites

To continue, the number of governmental sources is without any doubt high. This is not a random event as in a health crisis data from governmental sources and when they become available make it the first choice in the news' field. In a parallel line, they can perform better for crisis management such as happened in China for Covid-19. There, the data sharing, enhanced significantly the effort of dealing with the pandemic and raised the crowds' faith in the protection measures as well as in the Chinese government on this topic (Li et al. 2022; Xiao et al. 2021).

The Dataset "Govgr_Covid19_Misinfo"

With the same context, in April 2020, in the formal website of the Greek Ministry of Health for Covid-19, noticed a page addition relative to misinformation about the pandemic. This page called "Myths for Covid-19" and the main purpose of it is to gather all the misinformation sentences about Covid-19 and for each of them to give the appropriate truth answer. To succeed this, each truth sentence was given by a member of the Greek Committee of Infectious Diseases.


Myths about Covid-19 – Source: covid19.gov.gr

From this picture, it is understood that information management, data analysis and the extraction of reliable and useful statistical results, is difficult. Based on this, it is presented the "Govgr_Covid19_Misinfo" dataset. This dataset (.xlsx mode) is the structured and easily manageable format of the above page. It contains the following variables:

1. id: Id number

2. myths: The misinformation sentences

3. subject: The subject of each sentence (e.x. vaccine, antibiotics, weather conditions etc.)

4. aim: The main target of each sentence (e.x. protection - treatment from the virus, vaccine efficiency

5. date: Date of recording

6. vacc_timing: Timing relative to vaccine discovery

7. truth: The truth sentence of each incorrect sentence. It is worth to be noticed that all the truth sentences were written by a member of the Greek Committee of Infectious Disease.

You can download the "Govgr_Covid19_Misinfo" dataset in Greek or English from the following link: https://rb.gy/a9ypn0

# Data Analysis based on "Govgr Covid19 Misinfo" Dataset

In the previous part of this report, the dataset "Govgr_Covid19_Misinfo" presented in detail. Now, it is logical to provide a small data analysis based on this dataset. This analysis aims to present interesting insights. Firstly, the below graph presents the distribution of misinformation sentences' records about Covid-19 in a chronological sequence.
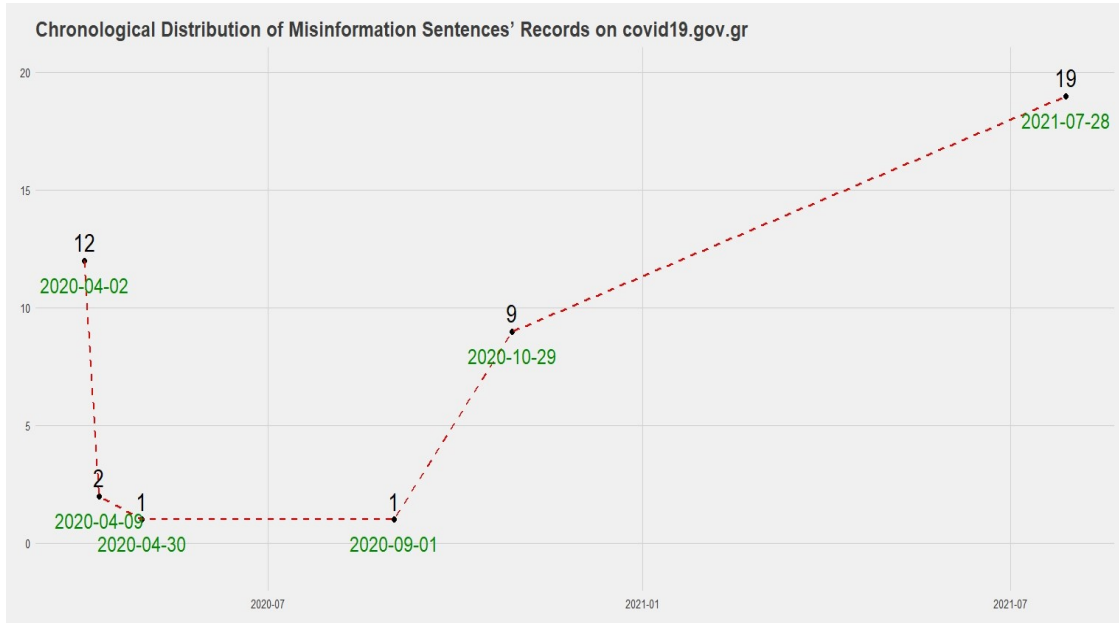
*Chronological Distribution of Misinformation Sentences' Records on covid19.gov.gr*

Based on this graph a technical result is extracted, that of the non-automatically update of this dataset. This choice could be explained by the fact that for each false sentence a true one is provided by a scientist, and this action is impossible to be successful in a dynamic way.

The next graph shows that the majority of the sentences have as subject the "Vaccine" while the other sentences have different subjects such as "Disinfectants and Antiseptics" and "Ultraviolet Lamps (UV)".
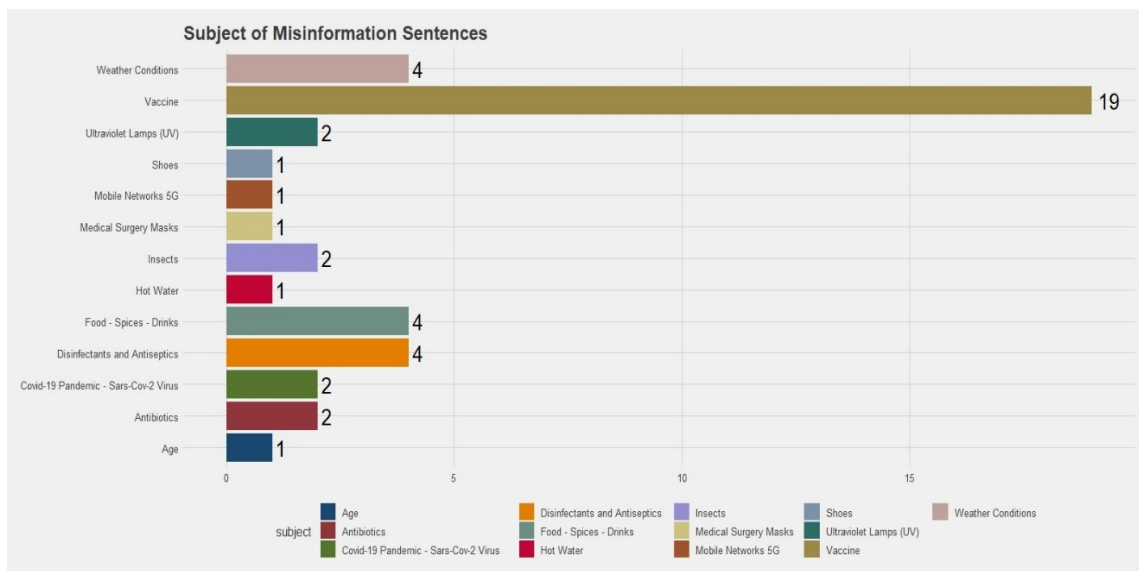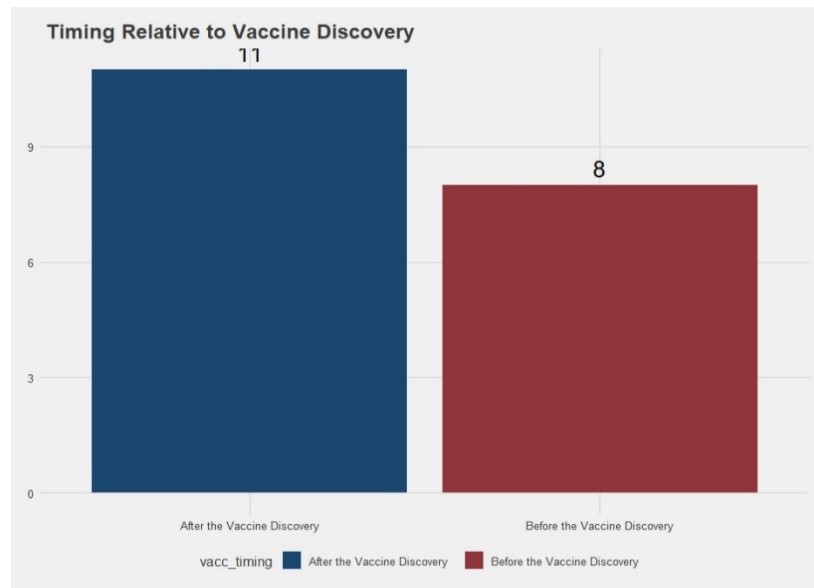

Subject of Misinformation Sentences

To continue, the next graph aims to present the relations between the misinformation sentences and the Covid-19 issues or problems of daily life. So, 18 sentences are related to "Protection – Treatment from the virus" and 4 sentences to "Transmission of Virus" while other sentences indicates topics like "Vaccine Efficiency", "Women Fertility" and "Nostosis from the Sars-Cov-2 virus".
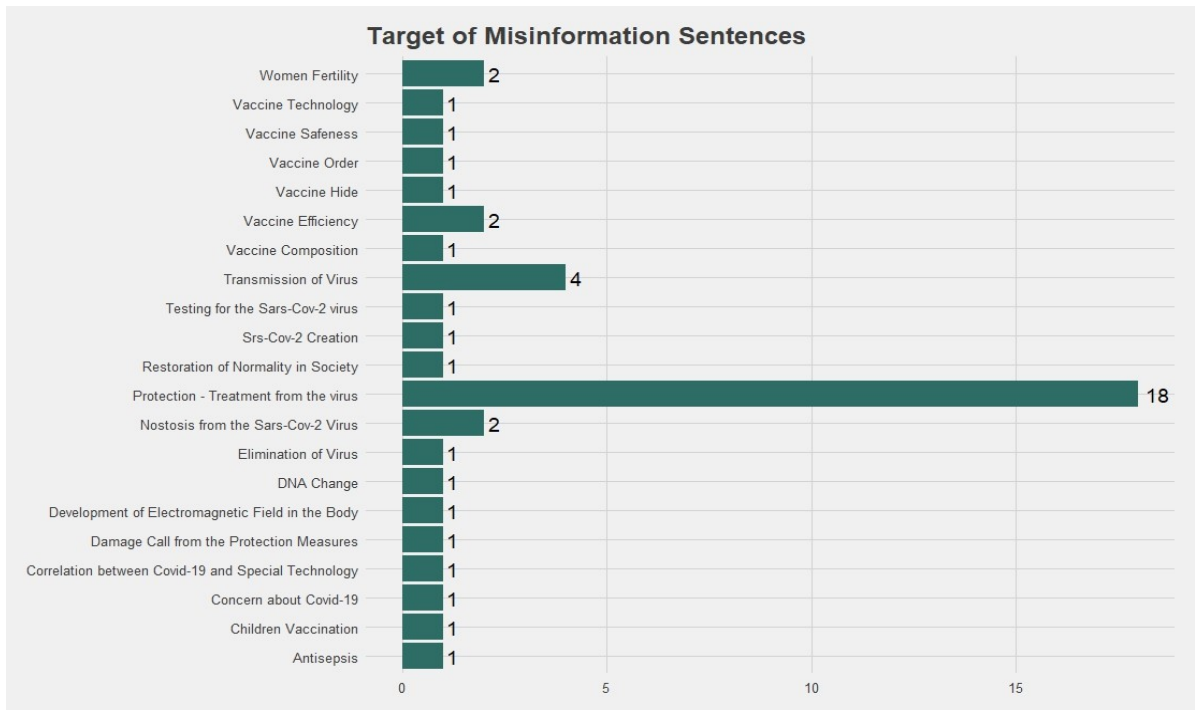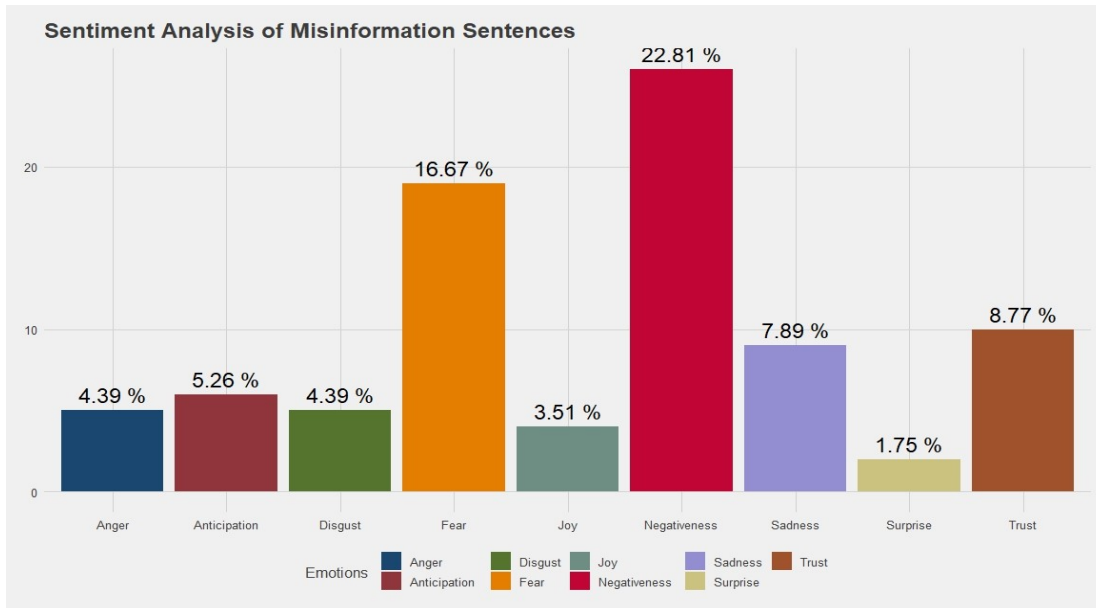

Target of Misinformation Sentences

According to the following graph, 19 misinformation sentences are linked to Covid-19 Vaccine. From them, 11 sentences dated before the vaccines' discovery while 8 dated after the vaccines' discovery.


Timing Relative to Covid-19 Vaccines' Discovery

Finally, the next 2 graphs focus on the sentiment analysis of misinformation sentences. From the first graph, is obvious to say that the 2 main emotions which extracted from these sentences are "Negative" (22.81%) and "Fear" (16.67%). Furthermore, the emotion of "Trust" has a percentage 8.77%, a not a negligible percent. The reason for this is that many sentences are beliefs for specific topics. A typical example is the sentence "When the vaccine is found, we will be fully restored to our old normalcy", because hides much trust to the vaccine while the normalcy's restoration is considered as certain.

Sentiment Analysis of Misinformation Sentences

The second graph is word-cloud of the sentences and clearly shows that the most common words were: "vaccine", "vaccines", "virus", "will", "new", "body", "kill" etc…
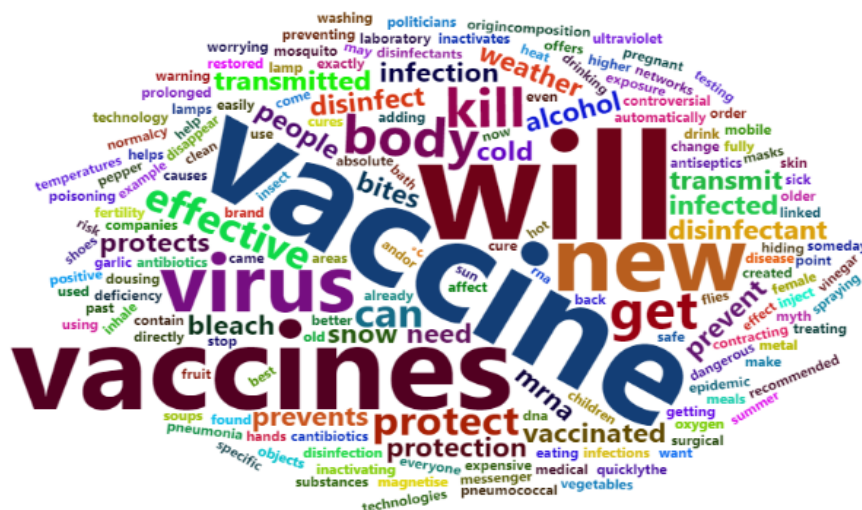
Word-Cloud Misinformation Sentences

Conclusion

Covid-19 Pandemic put the global community into a difficult situation. Despite the many problems in many systems (i.e. health systems), the pandemic indicated and the phenomenon of misinformation, too. To tackle the spread of it during the pandemic, decisions had been taken in a lot of countries as well as in Greece, with their effectiveness being judged either high or low. However, this timeless phenomenon must be managed for better and more correct people's information.

In general, governments and mass and social media must cooperate to take measures against the transmission of false news, without these measures being the product of special political interests and expediency. Finally, everyone should filter and cross-check the information he or she receives, while making the most of his or her critical faculties and trusting the unbiased opinions of the scientific community, especially on problems that threaten human life, such as that of a health crisis.

References

Anwar, Ayesha, Meryem Malik, Vaneeza Raees, and Anjum Anwar. 2020. "Role of Mass Media and Public Health Communications in the COVID-19 Pandemic." *Cureus*, September. https://doi.org/10.7759/cureus.10453.

Aspriadis, Neofytos. 2021. "Managing COVID-19 Pandemic Crisis: The Case of Greece." *Journal of International Crisis and Risk Communication Research* 4 (2): 387–412. https://doi.org/10.30658/jicrcr.4.2.8.

Brennen, J. Scott, Felix M. Simon, Philip N. Howard, and Rasmus Kleis Nielsen. 2020. "Types, Sources, and Claims of Covid-19 Misinformation." https://doi.org/10.60625/RISJ-AWVQ-SR55.

Giannouli, Iliana, Achilleas Karadimitriou, Ioanna Archontaki, and Stylianos Papathanassopoulos. 2024. "COVID-19 Vaccine Hesitancy: Analyzing Anti-Vaccine Rhetoric on Greek Facebook." *Online Media and Global Communication* 3 (2): 235–60. https://doi.org/10.1515/omgc-2024-0008.

Gisondi, Michael A, Rachel Barber, Jemery Samuel Faust, Ali Raja, Matthew C Strehlow, Lauren M Westafer, and Michael Gottlieb. 2022. "A Deadly Infodemic: Social Media and the Power of COVID-19 Misinformation." *Journal of Medical Internet Research* 24 (2): e35552. https://doi.org/10.2196/35552.

Li, Zhigang, Manjia Wang, Jialong Zhong, and Yiling Ren. 2022. "Improving the Communication and Credibility of Government Media in Response to Public Health Emergencies: Analysis of Tweets from the WeChat Official Accounts of 10 Chinese Health Commissioners." *Frontiers in Public Health* 10 (July). https://doi.org/10.3389/fpubh.2022.900776.

Liao, Tim F. 2022. "Understanding Anti-COVID-19 Vaccination Protest Slogans in the US." *Frontiers in Communication* 7 (June). https://doi.org/10.3389/fcomm.2022.941872.

Madraki, Golshan, Isabella Grasso, Jacqueline M. Otala, Yu Liu, and Jeanna Matthews. 2021. "Characterizing and Comparing COVID-19 Misinformation Across Languages, Countries and Platforms." *Companion Proceedings of the Web Conference 2021*, April, 213–23. https://doi.org/10.1145/3442442.3452304.

Marr, Bernard. 2018. "How Much Data Do We Create Every Day? The Mind-Blowing Stats Everyone Should Read." *Forbes* 21: 1–5.

Nsoesie, Elaine Okanyene, Nina Cesare, Martin Müller, and Al Ozonoff. 2020. "COVID-19 Misinformation Spread in Eight Countries: Exponential Growth Modeling Study." *Journal of Medical Internet Research* 22 (12): e24425. https://doi.org/10.2196/24425.

Organization, World Health et al. 2021. "Coronavirus Disease (COVID-19) Advice for the Public: Mythbusters."

Raballo, Andrea, Michele Poletti, and Antonio Preti. 2022. "Vaccine Hesitancy, Anti-Vax, COVID-Conspirationism: From Subcultural Convergence to Public Health and Bioethical Problems." *Frontiers in Public Health* 10 (May). https://doi.org/10.3389/fpubh.2022.877490.

Soranno, Patricia A., Edward G. Bissell, Kendra S. Cheruvelil, Samuel T. Christel, Sarah M. Collins, C. Emi Fergus, Christopher T. Filstrup, et al. 2015. "Building a Multi-Scaled Geospatial Temporal Ecology Database from Disparate Data Sources: Fostering Open Science and Data Reuse." *GigaScience* 4 (1). https://doi.org/10.1186/s13742-015-0067-4.

Wardle, Claire, and Hossein Derakhshan. 2017. *Information Disorder: Toward an Interdisciplinary Framework for Research and Policymaking*. Vol. 27. Council of Europe Strasbourg.

Xiao, Yao, Wanting Xu, Shouzhen Zeng, and Qiao Peng. 2021. "Online User Information Sharing and Government Pandemic Prevention and Control Strategies-Based on Evolutionary Game Model." *Frontiers in Public Health* 9 (November). https://doi.org/10.3389/fpubh.2021.747239.