REVISTA ESPAÑOLA DE COMUNICACIÓN EN SALUD
2024, v. 15, n. 2, 56-73
https://doi.org/10.20318/recs.2024.8186

Originales

Analysis of the social discourse on COVID-19 in the Spanish digital media

Análisis del discurso social sobre la COVID-19 en los medios digitales españoles

Begoña Rivas Rebaque[a], Manuel Gertrudix[b], Andres Chihuailaf[c], Alejandro Carbonell-Alcocer[b], Juan Romero-Luis[a], Alberto Sanchez-Acedo[a]

[a] Grupo de investigación Ciberimaginario, Universidad Rey Juan Carlos, España
[b] Departamento de Comunicación Audiovisual y Publicidad, Universidad Rey Juan Carlos, España
[c] Director Spain/Italy in Digimind, Francia

Abstract

Introduction: The media and social media have informed the public about the details and the evolution of the COVID-19 pandemic and the subsequent prevalence of Long COVID. There has been growing concern throughout that information should be accurate, sufficient, and truthful. **Objective:** The aim of this research is to identify this discourse in the Spanish digital media world in order to understand the media representations they produce and to improve communication awareness campaigns. **Methods:** Adopting a threefold methodological approach (Social Medial Listening, discourse analysis and social representation), the media discourse of Long COVID on different social media and in Spanish digital media is analysed for 12 months using the DIGIMIND tool. The five components of Laswell's Model of Communication: who; (says) what; (to) whom, which (channel) and what (effect), are evaluated and reported. **Results:** The analysis shows, among other results, which keywords are most present in social media and in digital press; the most consulted sources of information; or how information links are established between those studied. **Conclusions:** Based on these findings, an effective communication guide on Long COVID has been developed with a set of recommendations to be implemented by government agencies, health communication specialists and affected groups.

Keywords: Coronavirus; Digimind Software; Discourse Analysis; Health Communication; Long-COVID; Mass media.

Resumen

Introducción: Los medios de comunicación y las redes sociales han informado al público sobre las circunstancias y evolución de la pandemia del COVID-19 y la posterior incidencia del COVID persistente. En este proceso ha crecido la preocupación porque la información sea precisa, adecuada y veraz. **Objetivo**: caracterizar este discurso en el ecosistema mediático digital español para comprender las representaciones mediáticas que producen y mejorar las campañas de concienciación en comunicación. **Metodología**: Adoptando un enfoque triple (escucha en redes sociales, análisis del discurso y representación social), se analiza el discurso mediático sobre COVID persistente en diferentes redes sociales y en medios digitales españoles durante 12 meses utilizando la herramienta DIGIMIND. Se evalúan los cinco componentes del Modelo de Comunicación de Laswell: quién; (dice) qué; (a) quién, a través de (qué canal) y con qué (efecto). **Resultados**: El análisis muestra, entre otros, qué palabras clave están más presentes en las redes sociales y en prensa digital, las fuentes de información más consultadas, y cómo se establecen vínculos informativos entre los actores estudiados. **Conclusión**: Basándose en estos hallazgos, se ha desarrollado un decálogo de recomendaciones sobre COVID persistente para ser aplicadas por organismos oficiales, especialistas de comunicación en salud y colectivos de afectados.

Palabras clave: Análisis del discurso; Comunicación en salud; Coronavirus; COVID persistente; Medios de comunicación; Software Digimind.

CONTACTO: Begoña Rivas Rebaque begona.rivas@urjc.es

Fecha de envío:09/11/2023
Fecha de aceptación: 08/07/2024

Introduction

In March 2020, the Director-General of the World Health Organization (WHO) announced that COVID-19 should be considered a global pandemic and stated "this is not just a public health crisis, it is a crisis that will affect all sectors, and for that reason all sectors and everyone must take part in the fight" (WHO, 2020a). However, beforehand (4th January 2020), this non-governmental organization reported a cluster of non-fatal cases of pneumonia in Wuhan on its Twitter profile and announced that investigations were underway to specify the cause of the disease (WHO, 2020b).

In this regard, the pandemic has fostered a widespread use of technology, connectivity and time spent interacting on social media using various devices, globally (IAB, 2022; Fundación Telefónica, 2022; Rivero, 2021). For example, 97.1% of the world's population has a mobile phone and the time spent by internet users on social media on any device is 2 hours and 25 minutes (Rivero, 2021). The fact that this type of news is communicated through social media is significant given the characteristics of social media: immediacy; frequency of use among internet users; and the reach of some profiles such as the WHO's Twitter, which exceeded 11.9 million followers in December 2022.

The IAB Spain 2022 study on social media indicates that 85% of Spanish internet users aged 12 to 70 use social media, 51% of whom are women

Several reports have been produced in Spain, the results of which provide relevant contributions on the exponential growth, consumption habits and uses of technology, motivated, in part, by the need to use digital, social and mobile channels in our daily activity during lockdown. Thus, the Mobile Report 2021 –in Spain and the World (Rivero, 2021)– highlights that 95% of Spaniards have messaging and social networking applications installed on their mobile phones. The IAB Spain 2022 study on social media indicates that 85% of Spanish internet users aged 12 to 70 use social media, 51% of whom are women. This report also indicates that the main reasons for their use include entertainment, interaction and information, and that Gen-Z is the generation that uses social media the most, with an average of 6.3 networks; and influencers' credibility is 47%.

Fundación Telefónica's Digital Society in Spain Report 2022 refers to the use of social media as a means of accessing digital and audiovisual content (2022), and the report on Benefits and Risks of Internet and Social Network Use 2022 (ONTSI, 2022a) suggests that "social networks have become increasingly important as a channel for connecting [...] with others" (2022, p.8). In addition, respondents believe "that factors such as time spent and frequency of use, enhance the undesired effects, especially on social media", (2022, p.43), and attribute greater responsibility for the prevention of this inappropriate use to the companies that manage social networks, after the Public Administration. At the same time, they mention the work of the media, which it can favors the responsible use of technologies, and believe that specific legislation on the use of social media is necessary (2022).

The National Observatory of Technology and Society (ONTSI, 2022b) emphasises that social networks are a great source of data for companies. In addition, the Networked Society 2019 annual report (ONTSI, 2020) specifies that social media is the main social measure used by companies, and, in turn, warns of the risks that countries should identify and address, including those related to the sharing of information and the subsequent use made of it, which could lead to manipulation of social media, among other concerns. As demonstrated by Bae et al. (2021) the influence of misinformation on social media also affects the predictive accuracy of infection models, such as Susceptible-Exposed-Infectious-Recovered (SEIR), hence the importance of understanding diffusion models to use social and informational motivation strategies.

Background

Infodemics as a Communicative Framework

From a conceptual perspective, the study presented here is consistent with the Communitiy Guide´s theoretical definition of health communication (2022) as "the study and use of communication strategies to inform and influence individual and community decisions that enhace health", as well as the variety of communication channels available to deliver messages to different audiences, health professionals, individuals, communities, and decision makers. This communication initiative is also in line with the approach taken by Robert Hornik "The first step in developing an effective health communication strategy is to understand the targeted behavior" (IOM, 2015, para.1).

In addition, on a more applied level, this study ties into some of the suggestions and ideas contained in Culturally Appropriate Health Care, such as Kreps' idea that "The need for effective communication about health promotion is particularly acute in modern society, where there are significant health risks." (2007, p. 113) and the different types of recommendations, such as those mentioned by Betancourt (2007, p. 26) on eliminating racial/ethnic disparities in health care with a focus on a broad range of stakeholders, the health literacy strategies and interventions described by Levin-Zamir (2007, p. 109), or similar resources such as the booklet of recommendations on nutritional adaptation for health professionals explained by Jaffe et al. (2007, p. 121).

As the recommendation guide is a communication material, the principles for effective communication proposed by the WHO (n.d.) (accessible, actionable, credible, relevant, timely, and understandable) have been integrated to achieve better health outcomes and to address different audiences such as international organizations, policy makers, individuals, communities, and health care providers.

In this context, authors such as Mheidl & Fares (2020) revisit the basics of health communication during the pandemic COVID-19 and propose a "checklist for responding to the infodemic," Basch et al. (2020) assess the readability of information about coronavirus disease published on the Internet, and Allahverdipour (2020) emphasizes that "[...] only by working with communities and citizens, providing guidance and accurate information through reliable channels for public engagement, can the effectiveness of efforts against COVID -19 be ensured." Stolow et al. (2020) recommend the use of evidence-based health communication and provide support for health professionals in formulating health communication messages during COVID-19, while advising against the use of fear appeals.

In the context of the inappropriate use of digital behavior, the possible manipulation of information and its alleged consequences, it is important to point out that the Joint Statement carried out by the WHO and other International Organizations (WHO, 2020c, para.1) reported that "coronavirus disease (COVID-19) is the first pandemic in history where technology and social media are being used on a large scale to help people stay safe, informed, productive and connected", and addressed the infodemia or information overload, which includes deliberate attempts to spread misinformation that can harm people's physical and mental health, among other things.

In this framework, the 3rd WHO Global Infodemic Management Conference (WHO, 2020d) identified the stakeholders to address infodemic-related challenges as: 1) The scientific and research community; 2) Health authorities; 3) Technology companies and social media platforms; 4) NGOs and civil society groups; 5) Media and journalism; and 6) UN agencies and multilateral organizations. In turn, it called for the need for collaboration to ensure that good-quality health information is available to allow people to make the right decisions when it comes to protecting their health.

In this sense, WHO emphasises the importance of young people being informed about COVID-19 and practicing safe Internet surfing when making decisions to protect their health and that of their family members. This is according to the international study on social media and COVID-19 conducted by Wunderman Thompson, the University of Melbourne, Pollfish, and WHO, which explores where exactly Gen-Z and Millennials look for information about COVID-19; who they trust as reliable sources; and their awareness of fake news. The study found that scientific content is considered worthy of sharing (43.9%), and that 59.1% of respondents have a high awareness of COVID-19 related fake news, but also that users feel indifferent when it comes to countering or ignoring it (Volkmer, 2021).

Wang & Chen (2022) examine discourses on COVID -19 vaccination by collecting tweets and using the Lantent Dirichlet Allocation (LDA) model and sentiment analysis. Similarly, the work of Gadzekpo et al. (2023), which captured the opinions of journalists in Ghana using questionnaires, concluded that the majority of respondents had a positive attitudes towards COVID-19 vaccines and sought information from official sources to inform their work. Similarly, the study by Guan et al. (2022) examines the effects of news fatigue and what makes people tired in the context of the COVID-19 pandemic. Or the study by Huang et al. (2023), which sought to understand how the differences in information between communication channels are related to information overload and subsequent processing.

The analysis of social discourse in terms of its communicative components

The ongoing impact of COVID-19 has spawned a new media discourse and social narrative, the study of which is valuable in understanding how to improve the information conveyed by scientific research to patients, patient advocacy groups, health professionals, and communicators specializing in health information. Due to the uncertainty about the causes, consequences, and treatment of this disease (Alwan, 2021), it is even more important to provide

accurate information and understand the role that both the media (Anwar et al., 2020) and information published on the Internet and social media (Daraz et al., 2019) play in disseminating information and creating a social discourse that is so important to society.

To this end, the study takes into account Harold Lasswell's (1948, p.216) model or paradigm on the communication process and the components involved in communicative acts –who, says what, through what channel, to whom and with what effects?– analysing this communication process adapted to the "bidirectional" digital environment, since the messages are addressed to a multitude of people, and users can share the message, make comments, contributions or opinions or exchange information in the different digital media, and with a certain intentionality (Flores et al., 2021, p.287).

Approach

In this context mediated by the pandemic, and on the basis of a social discourse analysis, the research aims to analyse how Long COVID is being communicated in Spanish digital media (digital media and social media in Spanish), to understand the media and discursive representations they produce, and to promote communication strategies that can have an effect on COVID patients, helping them make suitable and correct communication decisions.

Method

The purpose of this approach is to provide operational recommendations for effective COVID communication campaigns based on the characteristics of the target groups.

Thus, the methodological approach is organized into three techniques:

1. The method of social listening or social media Listening to collect information on what users say on social media about COVID-19 and Long COVID with the aim of understanding the behavior of who is talking; what is being talked about; and/or who are the most relevant individuals that generate the most information on this topic. Social Media Listening is a research methodology widely used in Public Health studies, particularly in the studies on COVID-19 by Arillotta et al. (2021), Purnat et al., (2021), Jarynowski et al. (2021) and Ma et al. (2022). This methodology allows researchers to evaluate how digital social discourse on a specific subject is configured by analysing a large number of publications available on social media, digital press, websites, news, and blogs.

2. The analysis of media discourse in the digital media in order to understand the point of view of these media when generating and transmitting messages in different formats (written texts, videos, audios, etc.) and how they expose the reality of this issue when it comes to informing, suggesting, giving opinions, making claims, etc., in order to keep the public up to date and to be able to make the best decisions during this difficult period.

3. Social representation or people's common knowledge generated through experiences, beliefs and information given that, during the pandemic, social media and the digital press were used as a means of consultation, information, advice and even to feed the infodemic, defined by the WHO as "an overabundance of information, which may or may not be correct, during a pandemic" (WHO, 2020c).

Design

Considering the main objective of offering information of public interest, it should be noted that the social intelligence and social listening software of the company Digimind (now, Onclusive Social) was used for the collection of information. With this tool, which uses keyword-based source filtering, we were able to analyse the following variables using the various options available:

1. Which terms are most frequently searched for by internet users (Top reputation option).
2. What are the most cited keywords and hashtags (What option).
3. What are the social metrics in digital media (When option).
4. What are the most cited digital media (Where option).
5. What are the sources of information, their socio-demographic profiles and geographical locations (Who option).
6. What feelings are associated with the use of the analysed references (option How).

Similarly, for the data collection and analysis, the stages of the design process were taken into consideration in order to respond to the following research variables (Table 1), based on the communication components of Lasswell's model: information on who, what, which, whom, what, is provided.

Table 1. Relationship between Laswell' communication components and the research variables.

ID	Communication component	Variable	Description
V1	Who reports	Main contributors of the discourse	Who speaks
V2		Groups involved	Which groups are talking
V3	What is being reported	Content of the discourse	What is being spoken about
V4	When it is reported	Timing of the discourse	At what time the most interaction takes place
V5	Where it is reported	Channels or media	In which channels or media it is spoken
V6		Retrieval systems	What search engines are used
V7		Reference sources	What are the most consulted sources of information or references on the subject matter under investigation
V8	How it is reported	Sentiment analysis	What is the sentiment with which it is spoken (connotation of the language used)?

Source: Own elaboration.

The corpus of analysis is made up of publications in Spanish about COVID and Long COVID on social media (Facebook, Twitter, Instagram) and in the digital press (websites, news, blogs), over a period of 12 months (from 15/03/2021 to 15/03/2022). The period of one year was considered to coincide with the second anniversary of the first wave of the pandemic, from 15/03/2021 to 15/03/2022, with exceptions such as in the case of the analysis of the most searched terms (Top Reputation), where the week of March 2022 was taken as the time sample, given that the app cannot use data retroactive to the beginning of the search process. Likewise, for the analysis of keywords and information sources, this date range was extended to 05/09/2022 in order to obtain the most up-to-date sample possible of these variables to establish the guide for the treatment of Long COVID.

To the extent that information has been collected and data analysed on the variables that allow us to understand the informative treatment of Long COVID (see Materials section), the results have been communicated in the form of short reports or blog posts. These communications are part of the competitive research project LONG-COVID-EXP-CM, which was created expressly to disseminate information on this topic. At the same time, they have been shared via different social media platforms (Twitter, Facebook, Instagram, YouTube, TikTok) and content has been created in different multimedia formats (texts, infographics, interactive images, videos, ...) in order to offer greater communication and better understanding of the topic, and to be useful for the different groups targeted by this project.

The collection and analysis process carried out consists of the following stages:

Stage 1. Filtering sources by keywords in the Digimind app.
Stage 2. Harvesting information.
Stage 3. Data selection and cleaning.
Stage 4. Data analysis.
Stage 5. Creating a post.
Stage 6. Publishing a blog post.
Stage 7. Sharing the post on social networks.

Instruments

The following information was analysed using the Digimind tool:

What are the most searched terms by internet users? Top Reputation

"Top Reputation" data sheet

- Time period:14/03/2022 to 27/03/2022.
- Objective: To analyse which are the most popular results offered by Google to the internet user when searching for the keyword.
- Sample: Top 10 options suggested by the search engine.
- Methodology: Google Autocomplete Search: Monitoring and analysis/ Digimind.

Google's Autocomplete search system lists the most popular suggestions when internet users search for information about a term in the search engine. Thus, through the social monitoring platform Digimind, we analysed which are the most popular results offered by Google when searching for the keyword COVID, establishing the "Top Ten" results at a specific moment.

What is being talked about? What

"What" data sheet

- Time period: 15/03/2021 to 15/03/2022, and 15/03/2021 to 05/09/2022.
- Objective: To analyse which are the most cited keywords and hashtags in digital media and their subsequent categorization.
- Sample: 100 top hashtags and 35 key concepts.
- Methodology: Media discourse characterization study/ Digimind.

The social discourse in digital media on Long COVID increased as the incidence of COVID was observed. In this study on the characterization of media discourse in Spain, we analysed some of the keywords and hashtags most frequently cited in publications on different social networks such as Twitter, Facebook, Instagram and YouTube, as well as in other digital media such as news, blogs and websites.

When is the most interaction generated? When

"When" data sheet

- Time period: 15/03/2021 to 15/03/2022.
- Objective: To analyse what are the social metrics of COVID-19 and Long COVID in the digital media are.
- Sample: Mentions, interactions, estimated reach of messages on social networks (Twitter, Facebook, Instagram, YouTube), news, blog or website.
- Methodology: Trend analysis/ Digimind.

In this stage, an analysis of the prominence of the selected words was carried out in order to find out what their social metrics are, indicating the number of people talking; the amount of mentions; the estimated reach or the interactions that have been carried out for these hashtags according to the study's chosen time period.

Where is it being spoken about? Where

"Where" data sheet

- Time period: 15/03/2021 to 15/03/2022.
- Objective: To find out which are the most cited digital media on COVID-19 and Long COVID.
- Sample: Publications/posts about COVID-19 and Long COVID.
- Methodology: Breakdown of online channels/ Digimind.

An analysis was carried out to distinguish which media or channels generate the most mentions and the audience they have according to the studied topic.

Who is talking? Who

"Who" data sheet

- Time period: 15/03/2021 to 05/09/2022.
- Objective: To analyse the sources of information on COVID-19 and Long COVID, their socio-demographic profiles and geographical location.
- Sample: Top 100 influencers.
- Methodology: Filtering and score-based analysis of information sources and subsequent categorization/ Digimind.

The sources of information or those who make up the set of references, i.e., people (influencers), institutions and

media that contribute by generating and disseminating information on the COVID-19 and Long COVID issues were evaluated, since those who generate the discourse can have an impact or can influence the decision-making of any member of the public. In this respect, the Digimind tool made it possible to find out who are the people or networks of influencers who are talking about a given topic.

How is it spoken? How

"How" data sheet

- Time period: 15/03/2021 to 15/03/2022.
- Objective: To explore what are the sentiment connotations associated with the use of COVID-19 and Long COVID.
- Sample: Publications/comments/posts about COVID-19 and Long COVID.
- Methodology: Sentiment analysis/ Digimind.

The last stage focused on carrying out a sentiment analysis in order to recognize the positive, negative or neutral connotations associated with the use of references related to the topic under study. This makes it possible to assess the users' intentionality when posting the messages.

For the presentation and publication of the information, interactive applications such as Genial.ly or Flourish have also been used in order to improve the effectiveness of communication by means of visual presentation tools that bring the results closer to the users.

Results

The results are discussed to derive recommendations for improving health communication campaigns. The findings obtained for each of the study variables are specified below, grouped according to the communicative components; the period analysed; the intended objective; the sample studied; and the methodology applied.

Top Reputation

A filtering and search were carried out on the Digimind platform for this variable, in the specified time period. A primary analysis indicates that when internet users search for "COVID" in Spanish on Google, the first ten search engine positioning options are to do with the national or regional territory (Figure 1).

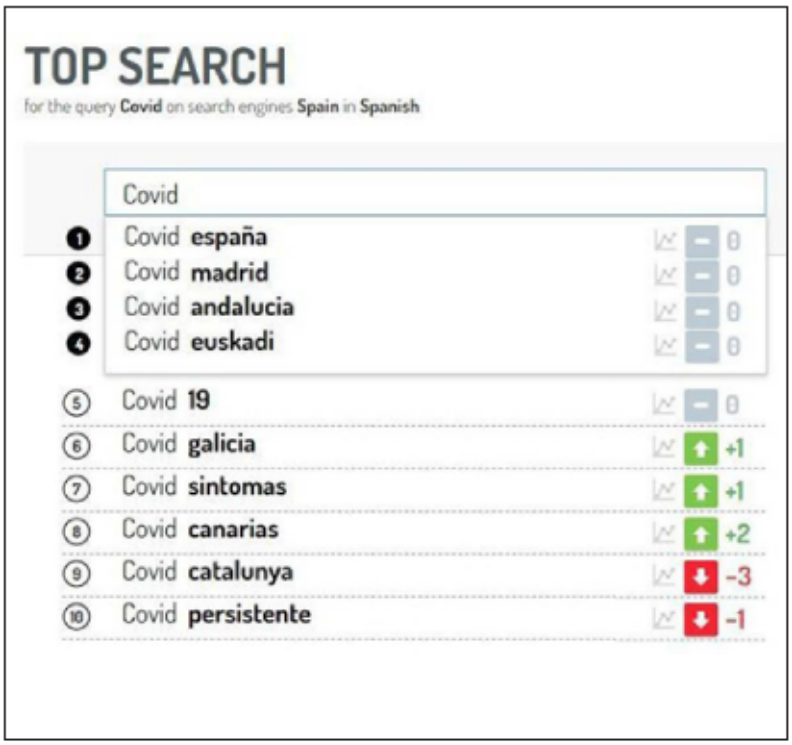

Figure 1. Top ten options suggested by Google when searching for the word "Covid". Source: Digimind Platform.

In order to find out which reference sources were consulted most frequently, the first query (Covid Spain) was analysed in more detail. The result shows that government sources were the most consulted, followed by media that provide information of an economic or statistical nature on the epidemiological situation (Figure 2). When analysing other similar queries, the results confirm these trends.

Figure 2. Reference sources of the search Covid España ("Covid Spain in English"). Source: Digimind Platform.

What

Two years after the pandemic was declared and given that the media serve to keep the public informed, a first analysis of the structure and composition of the first 100 hashtags studied was carried out. It is evident that they are in line with the subject matter which they are to do with and most of them are short (composed of a single word), simple and easy to understand and remember.

Similarly, the top ten most frequently cited hashtags were extracted, among which #covid19, followed by the compound words #Longcovid and its Spanish translation #covidpersistente, as well as some of their different forms of wording #covid and #covid_19, stood out in the top results (Figure 3).

about COVID 19 between Mar 15, 2021 and Mar 15, 2022

Rank	Hashtags	Mentions
1	#covid19	3 K
2	#longcovid	3 K
3	#covidpersistente	2 K
4	#covid	1 K
5	#coronavirus	767
6	#sarscov2	624
7	#covid_19	385
8	#pandemia	247
9	#salud	221
10	#covid-19	147

Figure 3. Top hashtags period 15th March 2021 to 15th March 15, 2022. Source: Digimind Platform.

Once these 100 hashtags were extracted, they were categorized by thematic categories according to the similarity of the words or because they shared some element related to the category to which they belonged. As a result, an infographic was created detailing the categories including the different hashtags, and additional information was added to each of these thematic categories for a better understanding of the content in which these characters were included, as well as some examples of their use. The thematic categories of the keywords obtained from Digimind´s platform are included in the following link: https://view.genial.ly/63f8e8d0b358c30012e64912.

Likewise, in order to understand the emerging media discourse that was taking place on social networks (Facebook, Twitter, Instagram) and in the digital press (websites, news, blogs), the presence of key concepts related to COVID-19 and Long COVID was analysed, extending the time frame from 15/03/2021 to 05/09/2022. The results presented in this interactive graph (https://public.flourish.studio/visualisation/12918762/) show the most used key concepts from 03/15/2021 to 09/05/2022, obtained from the data in Digimind´s platform, and clarify that the words "Persistent Covid" (in Spanish "Covid Persistente") in all its different forms of wording comes first place in mentions (10,991), followed by the English term "Longcovid" (as well as "Long Covid") with 3,179, and the Spanish words for "pandemic" (1,909), "disease" (1,058), and "side-effects" (1,032), among others.

When

The analysis of the prominence, or trending character of the hashtags (Figure 4), revealed that their peak occurred in the month of April 2021, which may coincide with the vaccinations for various age ranges that started during those dates. This is after the Ministry of Health, in collaboration with the Autonomous Communities, resumed the vaccination programme on 23rd March 2021 of the COVID-19 vaccination strategy. This result coincides with the growing presence of information in some digital media on the development of the vaccination, or headlines stating that Spain had beaten its record for the number of vaccines administered in one day.

Where

With regard to the analysis of the most cited digital media, this list of references to COVID-19 and Long COVID is led by the social media platform, Twitter, in terms of the place where most references were made (9,312 mentions); followed by other digital media or news (6,550 mentions), Facebook (2,666 mentions), blogs (265 mentions), and Instagram (98 mentions) (Figure 5).

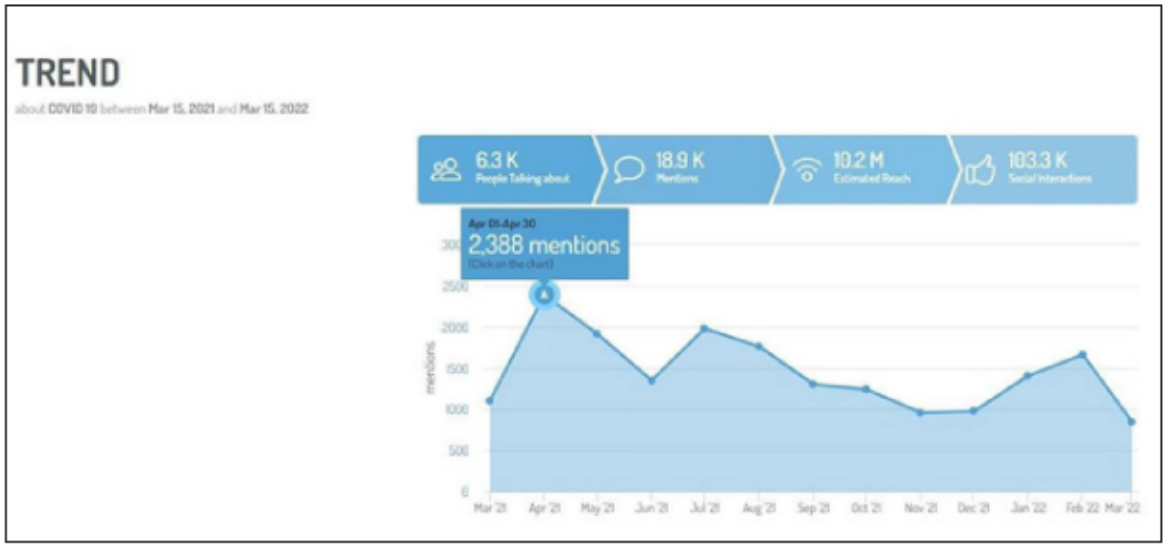

Figure 4. Trend analysis of references. Source: Digimind Platform.

Figure 5. Analysis of references in different media. Source: Digimind Platform.

Who

In order to find out who made up the group of people (influencers), institutions and media that were contributing to generating and disseminating information on the subject under study the most, the top 100 influencers were analysed according to the "score" of these sources. This indicator offers a score that establishes the influence of the source, obtained by multiplying its ranking by the number of mentions. Based on this data, the social profiles of the information sources were categorized into the following groups: Collective (25); Governmental (1); Media (21); Personal (53). The social profiles that generate and disseminate information about COVID-19 and Long COVID, obtained from Digimind´s platform, are linked in the next digital content: https://view.genial.ly/63fb4365fcda4600185c2b38.

The analysis of demographic data, such as gender and age of the people who gave their opinion and reported on this topic, showed that the majority of the audience was male (58% , compared to 42% female, and that most of the users were between 18 and 25 years old, especially women (Figure 6).

As for the geographical location of the main news sources analysed, as can be seen in the next interactive map (https://public.flourish.studio/visualisation/12913460/), the distribution is very heterogeneous throughout Spain, but there are some autonomous communities, such as Madrid, which concentrate a significant part of these news sources with the greatest impact.

Similarly, an analysis was carried out of the network graph of influencers or people who reported on COVID-19 and Long COVID that is represented in the following interactive chart: https://public.flourish.studio/visualisation/12913166/. This allowed us to find out how the information links between the people studied were established and how the structure of this network was configured. To this end, the centrality of the nodes was analysed, in which the size of the circle represents the number of interactions (re-tweets, quotes, etc.) or degree of influence. This included different groups such as SEMG, LongCovid ACTS, Yo Soy Long Covid, as well as the institutional account of the Ministry of Health, a personal account in the field of health, and a digital media outlet: elDiario.es.

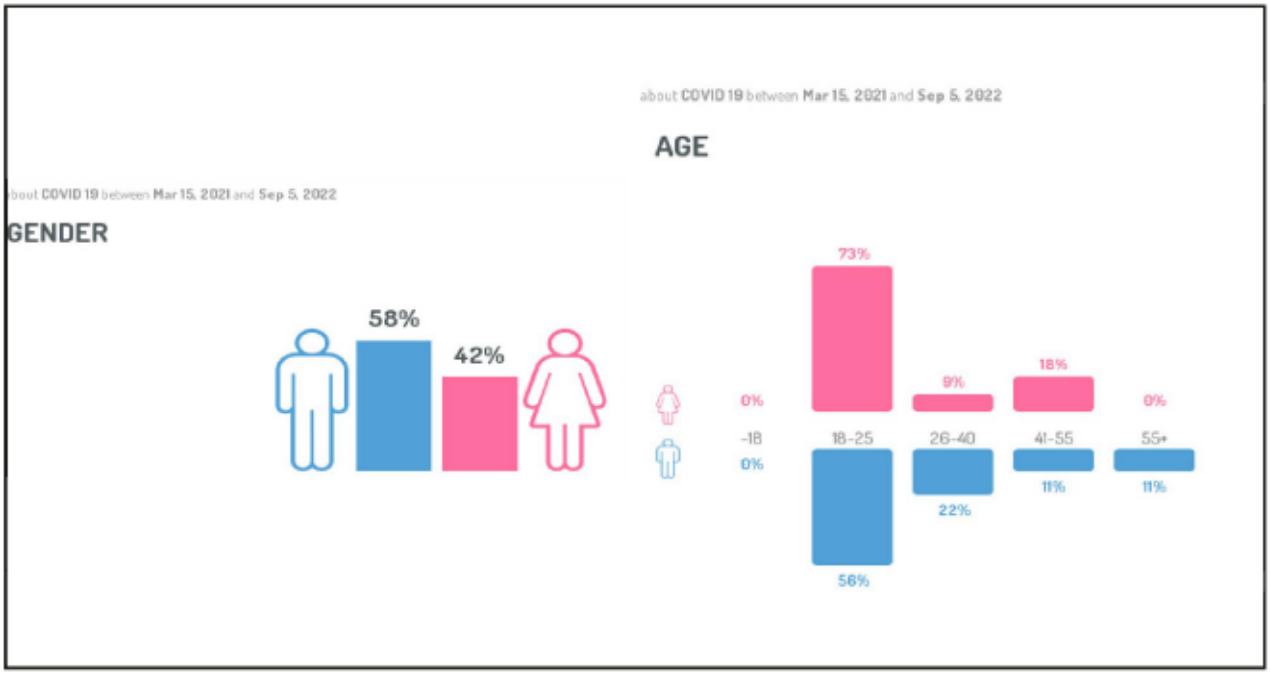

Figure 6. Socio-demographic profile of informants. Source: Digimind Platform

How

The sentiment analysis of the Digimind tool shows the positive, negative or neutral connotations associated with the use of the mentions under study, making it possible to evaluate the users' intentionality when using them. Thus, as shown in Figure 7, most of the posts carry a negative emotional tone (62% , being the most used in the opinions expressed in the different media analysed; followed by posts in which positive evaluations are shown (21% ; and expressions of neutral sentiment at the bottom (17% .

Figure 7. Sentiment analysis of mentions in posts. Source: Digimind Platform

Discussion and conclusions

Taking into consideration the preliminary results presented above, it can be concluded that the most consulted sources of information are those that correspond to reference and trustworthy sites such as, the Ministry of Health, the Ministry of Science and Innovation, the official information of the Spanish Government's COVID-19 vaccination strategy, the Web Portal for citizen access to information of interest of the State Administration and the National Institute of Statistics, etc.

Likewise, the analysis of the network of influencers or people who report on the topic shows that the number of followers of the different influencer networks is not directly proportional to the number of interactions carried out.

The analysis of the influencer network shows that the number of followers is not directly proportional to the number of interactions

The non-automated analysis of the posts shows that the communicative purpose of the texts and the intentionality in the construction of the media discourse generated by the effects of COVID-19 fulfils different functionalities: expressing opinions and/or supporting ideas; presenting information or describing facts or events; encouraging or promoting an action; advising, guiding, teaching, influencing or attracting attention; warning about facts or events; connecting with the receiver and exchanging information, and emotional communicative functions with the idea of expressing, telling, narrating experiences, feelings, states of mind, etc. Similarly, different communicative intentions can be observed implicit in the context of the same word depending on the message to be conveyed.

Thus, it is perceived that the purpose of providing data, research, indications, and clarifications on COVID-19 and Long COVID have enabled the public to make decisions to adapt and cope with the health crisis, as well as to meet their specific information needs an aspect that has been pointed out by Houben-Wilke et al. (2022)'s study; and to increase the ability to react to crises by Coombs (2020). This idea is further supported by health communication strategies, as they aim to change people's knowledge, attitudes, and/or behaviors, for example, to empower them to change or improve their health status (Rural Health Information Hub, 2023). In line with these ideas, it is worth noting the works of López Doblas et al. (2022) who analyse the use that Andalusian hospitals have made of social networks as the main communication channels in the management of the pandemic. The study by Villegas-Tripiana et al. (2020) investigates the resources published on the websites of public health institutions with information about COVID-19 for citizens. The research conducted by Ducci et al. (2022) investigates how Italian citizens have been informed about the pandemic, and the publication by Manrique-Grisales (2020) delves into the characters, institutions, territories, themes and lifestyles related to the pandemic through the contents of two media outlets from Ecuador and Colombia.

The results allow us to understand the extent to which a fake news story (as there are more producers of information nowadays) can go viral in an instant due to the greater reach of social networks, as Wineburg has concluded (Gragnani, 2018). It is therefore important that the information transmitted is reliable, accurate and truthful. In this regard, the study by Ma et al. (2022) discusses how different global health organizations communicate about the COVID-19 pandemic on social networks, and the strategies carried out, namely content prevention, information updating and communication guidance. And the research by Fernández-Zarza (2022) that carries out an analysis of the infodemic on COVID-19 in Spain in four organizations with the seal of the International Fact-Checking Network.

In this enormous flow of information, the communication of Long COVID does not always meet the needs and expectations of patients. Therefore, the ultimate aim of this study was to develop a communication guide (Gertrudix et al., 2022) to improve the informative treatment of Long COVID as a catalogue of advice/guidelines for institutions, such as patient associations and specialized health communication agencies, etc. This, in order to "enhance the safety, security and empowerment of individuals", is one of the digital rights and principles proposed in the EU for the Digital Decade (European Commission, 2022) and coincides with Cass Sunstein's premise that "democracies are based on a well-informed citizenry" (Gimeno, 2019).

In addition, it is necessary to emphasize consistency with some of the health communication considerations outlined in the Rural Health Information Hub, such as understanding the target audience to ensure the relevance of the content created and tailoring messages to the communication channel used to achieve greater reach (Rural Health Information Hub, 2023). Similarly, Kreps (2007) describes the need for audience analysis to guide the way messages are designed and delivered: "It is also important to assess consumers' language skills and orientations in terms of their culture, health literacy levels, motivations for seeking health information, and their particular patterns of media use" (2007, p.114). In this sense, Morales and Crespo (2022) state that the health sector understands that the patient "searches for information on the Internet and engine queries before select a pharmacy, a specialist or even when he does not know what happens and he look for people who are going through the same thing" (p. 63). Furthermore, this study shows that 50% of those affected by COVID-19 persistent (according to the SEMG survey) are between 35 and 50 years old, an "age in which social networks and digital are part of their daily life" (p. 64).

This decalogue of recommendations (https://view.genially.com/639e11589a40c90019f657da) has similar goals to the studies by Basch et al. (2020), Allahverdipour (2020), and Stolow et al. (2020) and coincides with some of the recommendations proposed by Levin-Zamir (2007) such as "Develop standards for appropriate messages and applying them in the development of materials using formative Research and make more efficient use of culturally appropriate mass media and education- entertainment" (2007, p. 109), in which it is organized into short slogans with concise explanations that address infodemics, intentionality and emotional tone of the message, how to achieve greater interaction and dissemination of publications, verification of data and citation of sources, etc.

Limitations and future research

In the analysis of the WHO variable, and to ensure appropriate use of the direct or indirect personal data displayed in the social profiles, although they were public profiles the nicknames (usernames @) were not cited in the examples used, and just looked at the profession detailed in the bio of these personal accounts. Thus, they were categorized into three groups: a) careers related to Medicine and Health, b) other careers and c) users concerned with and interested in COVID-19 and Long COVID.

Similarly, to specify the location of the different social profiles analysed, it should be noted that not all the accounts analysed included the location, therefore, the following should be added to the locations indicated on the map: 20 accounts that state Spain as the location (9 Group Accounts, 5 Personal Accounts, and 6 Media Accounts); 5 without specifying a location (3 Group Accounts and 2 Personal Accounts) and 1 Personal Account in which another country, in addition to Spain, is indicated.

Similarly, regarding the sentiment analysis, to properly interpret the result obtained, it should be noted that this is an analysis extracted automatically from the language used in the publications.

It should also be considered that the socio-demographic data analysed with Digimind includes the names of associations, platforms, government accounts, digital media, etc., as well as the names of personal profiles. In addition, there are other limitations derived from possible biases that may arise from the use of an automatic extraction model in the selection of sources and discourses (Hargittai, 2020) or in the presentation of certain social

groups (Pokhriyal et al, 2023). To address these limitations, further studies should be conducted to replicate the model adjusting for participation biases to ensure that social media analysis accurately reflect broader public opinion.

The results obtained are preliminary, so as future research we propose to extend the period of these same variables to find out 1) if the references studied continue to be a trend; and 2) to compare whether the sources of information and the digital media consulted in this new period coincide with those that have the information that those affected require; 3) to infer who is making the discourse on social networks, i.e. whether it is the users, the influencers or the media, and the type of connotations that continue to generate the media discourse on this topic.

Additional Information

Appendices

The Appendices can be found under [https://zenodo.org/record/7432212#.Y6IKt3bMKUl]

Disclosure Statement

No potential conflict of interest was reported by the authors.

Ethical Approval

The research was approved by the Research Ethics Committee of the Universidad Rey Juan Carlos on 27th June 2022 (internal registration number: 3005202214422).

Funding

This work was supported by the Autonomous Community of Madrid and the European Union (European Regional Development Fund-FEDER) [project number LONG-COVID-EXP-CM] and Universidad Rey Juan Carlos (ID 501100007511) with a grant call for Personnel in Training 2020 (PREDOC 20-008).

Acknowledgements

The analysis of media discourse has been carried out with the intelligence and social listening software of the company Digimind.

Author contributions statement

Conceptualisation: BR, MG; Methodology: BR, MG, AC, ACA; Formal analysis: BR, MG; Investigation: BR, ACA, JR, AS; Writing - Original Draft: BR, MG, ACA, JR, AS; Writing - Review & Editing: BR, MG, ACA; Visualization: BR; Supervision: MG; Project administration: MG; Software: AC; Resources: AC.

References

Alwan, N.A. (2021). The teachings of Long COVID. *Communications Medicine 1*(15):1-3. https://doi.org/10.1038/s43856-021-00016-0

Allahverdipour, H. (2020). Global challenge of health communication: infodemia in the coronavirus disease (COVID-19) pandemic. *Journal of Education and Community Health, 7*(2), 65-67. https://doi.org/10.29252/jech.7.2.65

Anwar A, Malik M, Raees V, Anwar A. (September 14, 2020) Role of Mass Media and Public Health Communications in the COVID-19 *Pandemic. Cureus* 12(9): e10453. doi:10.7759/cureus.10453

Arillotta, D., Guirguis, A., Corkery, J. M., Scherbaum, N., & Schifano, F. (2021). COVID-19 Pandemic Impact on Substance Misuse: A Social Media Listening, Mixed Method Analysis. *Brain sciences, 11*(7), 907. https://bit.ly/3FLu39P

Bae, S., Sung, E., & Kwon, O. (2021). Accounting for social media effects to improve the accuracy of infection models: combatting the COVID-19 pandemic and infodemic. *European Journal of Information Systems, 30*(3), 342-355. https://doi.org/10.1080/0960085X.2021.1890530

Basch, C. H., Mohlman, J., Hillyer, G. C., & Garcia, P. (2020). Public health communication in time of crisis: readability of on-line COVID-19 information. *Disaster medicine and public health preparedness, 14*(5), 635-637. https://doi.org/10.1017/dmp.2020.151

Betancourt, J. R. (2007) Racial and Ethnic Disparities in Health Care. *Epstein L, Culturally Appropriate Health Care by Culturally Competent Health Professionals. International Workshop Report. Caesarea. Israel: National Institute for Health Policy and Health Service Research* (pp. 18-32) https://acortar.link/6voKBG

Coombs, W. T. (2020). Public sector crises: Realizations from Covid-19 for Crisis Communication. *Partecipazione e conflitto, 13*(2). https://bit.ly/3I8Wqk3

Daraz, L., Morrow, A.S., Ponce, O.J., Beuschel, B., Farah, M.H., Katabi, A., Alsawas, M., Majzoub, A.M., Benkhadra, R., Seisa, M.O., Ding, J.F. (2019). Can patients trust online health information? A meta-narrative systematic review addressing the quality of health information on the internet. *Journal of general internal medicine 34*(9):1884-91. https://doi.org/10.1007/s11606-019-05109-0

Ducci, G., Boccia Artieri, G., Corsi, M., & Folena, C. (2022). Fuentes de información, confianza y comunicación del sector de la salud pública sobre la pandemia de Covid-19: Un estudio en la percepción de los ciudadanos italianos. *Revista Española De Comunicación En Salud, 13*(2), 182-199. https://doi.org/10.20318/recs.2022.7039

European Commission (2022, January 26). *Europe's Digital Decade: digital targets for 2030*. European Commission. https://bit.ly/3HS8Axw

Fernández Zarza, L. (2022). Análisis de la infodemia sobre la Covid-19 en España. *Revista española de comunicación en salud, 13*(2), 219-228. https://doi.org/10.20318/recs.2022.6410

Flores, A. E. H., Maldonado, A. T. C., & Sánchez, N. D. F. (2021). El modelo de Harold Lasswell aplicado en comunicación digital. Caso: La Posta. *Journal of Science and Research*, 6(3), 275-289. https://bit.ly/3UGVCFP

Fundación Telefónica (2022). *La Sociedad digital en España 2022*. (Report #SDIE 2022. ISBN: 978-84-306-2559-8) https://bit.ly/3WHDeNZ

Gadzekpo, A., Tietaah, G. K. M., Yeboah-Banin, A. A., & Kwame Ampofo Adjei, D. (2023). Media coverage of COVID-19 vaccines: sources of information, and verification practices of journalists in Ghana. *Journal of Communication in Healthcare, 17*(1), 15–29. https://doi.org/10.1080/17538068.2023.2208893

Gertrudix, M., Rivas-Rebaque, B., García Carrizo, J., Romero-Luis, J., Carbonell-Alcocer, A. & Sánchez-Acedo, A. (2022). *Guía para mejorar el tratamiento informativo del COVID persistente (Version 1)*. Zenodo. https://doi.org/10.5281/zenodo.7432212

Gimeno, R. (2019, February 12). Redes sociales *¿Qué le pasa a tu cerebro si dejas Facebook?* Retina tendencias. https://retinatendencias.com/

Gragnani, J. (2018, September 18). *Guía básica para identificar noticias falsas (antes de mandarlas a tus grupos de WhatsApp)*. BBC News Brasil. https://bbc.in/2MMnB3E

Guan, M., Li, Y., Scoles, J. D., & Zhu, Y. (2022). COVID-19 Message Fatigue: How Does It Predict Preventive Behavioral Intentions and What Types of Information are People Tired of Hearing About? *Health Communication, 38*(8), 1631–1640. https://doi.org/10.1080/10410236.2021.2023385

Guide to Community Preventive Services. (October 4, 2022). Health Communication and Health Information Technology. [The Community Guide]. https://bit.ly/42Jvf6K

Hargittai, E. (2020). Potential Biases in Big Data: Omitted Voices on Social Media. *Social Science Computer Review, 38*(1), 10-24. https://doi.org/10.1177/0894439318788322

Houben-Wilke, S., Goërtz, Y. M., Delbressine, J. M., Vaes, A. W., Meys, R., Machado, F. V., van Herck, M., Burtin, C., Posthuma, R., Franssen, F. M., Vijlbrief, H., Spies, Y., van 't Hul, A. J., Spruit, M. A., & Janssen, D. J. (2022). The Impact of Long COVID-19 on Mental Health: Observational 6-Month Follow-Up Study. *JMIR mental health*, *9*(2), e33704. https://doi.org/10.2196/33704

IAB Spain (Interactive Advertising Bureau) Spain. (2022). *Estudio de Redes Sociales 2022*. Elogia. https://bit.ly/3EfE6mQ

IOM (Institute of Medicine). 2015. Communicating to advance the public's health: Workshop summary. Washington, DC: The National Academies Press. https://bit.ly/3MmZfhF

Jarynowski, A., Kaczmar, K., & Madej, M. (2021). Listening to Twitter about Adverse Events of the Comirnaty COVID-19 Vaccine during first weeks of immunisation in Poland. *E-methodology*, *7*(7), 85-92. https://bit.ly/3HuNNQr

Jaffe, Anat., Guttman, Nurit. & Schuster, Michal. (2007). The Evolution of the Tene Briut Model – Developing an Intervention Program for the Ethiopian Immigrant Population in Israel and its Challenges and Implications. *Epstein L, Culturally Appropriate Health Care by Culturally Competent Health Professionals. International Workshop Report. Caesarea. Israel: National Institute for Health Policy and Health Service Research (pp. 121-142)*. https://acortar.link/calMzq

Kreps, G. (2007). Health communication at the population level: principles, methods and re-sults. *Epstein L, Culturally Appropriate Health Care by Culturally Competent Health Pro-fessionals. International Workshop Report. Caesarea. Israel: National Institute for Health Policy and Health Service Research* (pp. 112-120). https://bit.ly/42XcyMU

Lasswell, H. D. (1948). The structure and function of communication in society. In L. Bryson (Ed.), *The communication of ideas* (Vol. 37, No. 1, pp. 136-139).

Levin-Zamir, D. (2007). Health literacy, Empowerment and Cultural Competency. *Epstein L, Culturally Appropriate Health Care by Culturally Competent Health Professionals. International Workshop Report. Caesarea. Israel: National Institute for Health Policy and Health Service Research (pp. 101-111)* https://acortar.link/calMzq

López Doblas, M., Gómez Domínguez , D., Pérez Corral , O. R., & Danet Danet, A. (2022). Comunicar en tiempos de pandemia: Uso de las redes sociales y papel de las unidades de comunicación en cinco hospitales andaluces. *Revista española de comunicación en salud*, *13*(2), 155-172. https://doi.org/10.20318/recs.2022.6630

Ma, L., Wang, Y., & Kim, J. (2022). How health organizations communicate about COVID-19 on social media: a comparative content analysis. *Journal of Communication in Healthcare*, 1-10. https://bit.ly/3HRYwED

Manrique-Grisales, J. (2020). La prensa en modo pandemia: una aproximación a las agendas informativas de Expreso y El País sobre COVID-19. *Revista española de comunicación en salud*, 142-157. https://doi.org/10.20318/recs.2020.5450

Mheidly, N., & Fares, J. (2020). Leveraging media and health communication strategies to overcome the COVID-19 infodemic. *Journal of public health policy*, *41*(4), 410-420. https://doi.org/10.1057/s41271-020-00247-w

Morales Lorente, S. & Crespo, D. (2022). Comunicación entre personas afectadas por COVID-19 persistente con excelentes resultados en la visibilidad de la nueva enfermedad. *Revista de Ciencias de la Comunicación e Información*, 27, 52-66. https://doi.org/10.35742/rcci.2022.27.e271

Observatorio Nacional de Tecnología y Sociedad (ONTSI) (2022a). *Beneficios y riesgos del uso de Internet y las redes sociales*. 2022. Madrid. Ministerio de Asuntos Económicos y Transformación Digital. https://bit.ly/3UGkmOt

Observatorio Nacional de Tecnología y Sociedad (ONTSI) (2022b). *Uso de tecnologías digitales por empresas en España. 2022*. Madrid. Ministerio de Asuntos Económicos y Transformación Digital. https://doi.org/10.30923/094-22-006-6

Observatorio Nacional de Tecnología y Sociedad (ONTSI) (2020). *La Sociedad en Red. Transformación digital en España. Informe Anual 2019*. Edición 2020. Madrid. Madrid: Secretaría General Técnica, Centro de Publicaciones. http://doi.org/10.30923/1989-7424-2020

Organización Mundial de la Salud (OMS) (2020, March 11). *Alocución de apertura del Director General de la OMS en la rueda de prensa sobre la COVID-19 celebrada el 11 de marzo de 2020*. OMS web. https://bit.ly/3EcOdHZ

Organización Mundial de la Salud (OMS) (2020, September 23). *Gestión de la infodemia sobre la COVID-19: Promover comportamientos saludables y mitigar los daños derivados de la información incorrecta y falsa*. OMS web. https://bit.ly/3NTGgM7

Pokhriyal, N., Valentino, B.A. & Vosoughi, S. (2023) Quantifying participation biases on social media. *EPJ Data Sci.* 12, 26 https://doi.org/10.1140/epjds/s13688-023-00405-6

Purnat, T. D., Vacca, P., Czerniak, C., Ball, S., Burzo, S., Zecchin, T., ... & Nguyen, T. (2021). Infodemic signal detection during the COVID-19 pandemic: development of a methodology for identifying potential information voids in online conversations. *JMIR infodemiology, 1*(1), e30971. https://doi.org/10.2196/30971

Huang, Q., Mao, B., Jia, X., & Peng, W. (2023). COVID-19 Information Overload Mediated the Effects of Cross-Channel Information Differences on Health Information Elaboration. *Journal of Health Communication, 28*(7), 401–411. https://doi.org/10.1080/10810730.2023.2217097

Rivero, F. (2021). *Informe Mobile 2021 - España y Mundo*. Ditendria. https://bit.ly/3tjpX1H

Rural Health Information Hub. (2023). Health Communication. [Ruralhealthinfo.org] https://bit.ly/45l3FhZ

Stolow, J. A., Moses, L. M., Lederer, A. M., & Carter, R. (2020). How fear appeal approaches in COVID-19 health communication may be harming the global community. *Health Education & Behavior, 47*(4), 531-535. https://doi.org/10.1177/1090198120935073

Villegas-Tripiana, I., Villalba-Diaz, A., & López-Villegas, A. (2020). Análisis de la información sobre COVID-19 en sitios web de organizaciones públicas sanitarias. *Revista Española de Comunicación en Salud*, 234-242. https://doi.org/10.20318/recs.2020.5437

Volkmer, I. (2021). Social media and COVID-19: *A global study of digital crisis interaction among Gen Z and millennials*. World health organisation. https://bit.ly/3Wa3EZ6

World Health Organization (WHO) (2020a, March 11). *Alocución de apertura del Director General de la OMS en la rueda de prensa sobre la COVID-19 celebrada el 11 de marzo de 2020*. [WHO Web]. https://bit.ly/3EcOdHZ

World Health Organization (WHO) [@WHO] (2020b, January 4) *#China has reported to WHO a cluster of #pneumonia cases —with no deaths— in Wuhan, Hubei Province. Investigations are underway to identify the cause of this illness*. [Tweet] Twitter page. https://bit.ly/3tgTBEC

World Health Organization (WHO) (2020c, September 23). *Gestión de la infodemia sobre la COVID-19: Promover comportamientos saludables y mitigar los daños derivados de la información incorrecta y falsa*. OMS web. https://bit.ly/3NTGgM7

World Health Organization. (October-December 2020d). *WHO third global infodemic management conference: whole-of-society challenges and approaches to respond to infodemics*. Geneva: World Health Organization; 2021. ISBN 978-92-4-003450-1. Licence: CC BY-NC-SA 3.0 IGO. https://bit.ly/3Um0UHe

World Health Organization. (n.d.). *WHO principles for effective communications*. [Who.int] https://bit.ly/45grdEx

Wang, Y., & Chen, Y. (2022). Characterizing discourses about COVID-19 vaccines on Twitter: a topic modeling and sentiment analysis approach. *Journal of Communication in Healthcare*, *16*(1), 103–112. https://doi.org/10.1080/17538068.2022.2054196