The 11th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2020)
November 2-5, 2020, Madeira, Portugal

Mining text on Coronavirus

Cristian O. Beltrán-Oicatá[a]*, Mercedes Gaitán-Angulo[a], Melissa Lis-Gutiérrez[b], Alexa Senior Naveda[c]

[a]Fundación Universitaria Konrad Lorez, Bogotá, Colombia
[b]Corporación Universitaria del Meta, Villavicencio, Colombia
[c]Universidad de la Costa, Barranquilla, Colombia

Abstract

This document describes the behavior of the production of academic papers and high circulation newspapers, published under the terms "Coronavirus, Covid-19 or SARS-CoV-2". The purpose of this paper is to employ text mining techniques and qualitative data analysis for pattern identification (Dobrzyński, Dziekoński, Krzysztof, & Jurczuk, 2013; Westergaard, Staerfeldt, Tønsberg, Jensen, & Brunak, 2018). The interest of this study arises from the importance of this topic today and its relationship with the economy, business and society.

As a result, the production of scientific papers has been led by the United States, China and Russia. In Latin America, Brazil is the country with the highest scientific output on this subject. The analysis of newspaper articles focusing on economic and business issues showed the work between large companies, the government and the community, as well as companies with businesses and the need to adapt to constant change in the face of the crisis of labor and production.

Keywords: Coronavirus; SARS-CoV-2; COVID-19; pandemic.

* Corresponding author. Tel.: +57 319 2600844
 E-mail address: cristiano.beltrano@konradlorenz.edu.co

1. Introduction

Importantly, Covid-19 places consumers and businesses in continuous uncertainty on the verge of an unprecedented crisis. This crisis is affecting companies, workers, countries, economies, demand, supply, financial markets, at the global and local level, resulting in credit crises in sectors such as commerce, tourism, among others [1]

Governments have had to implement measures to control or reduce the contagion among their citizens, such as the closure of businesses and offices, the ban on travels that are not strictly necessary, even a mandatory quarantine has been imposed at home [2]. All of the above indicates that exports will be drastically reduced, resulting in changes in the GDP of the destination country and negative impacts on its economy

Analyzing the Special Report No. 3 "Social challenge in Times of Covid-19" issued by ECLAC, the following elements are found:

- Quarantine and physical distancing measures, necessary to curb the accelerated spread of coronavirus and save lives, generate job losses (in 2020, there would be 11.6 million more unemployed than in 2019) and reduce the labor income of individuals and households.
- The loss of income mainly affects the broad strata of the population living in poverty and vulnerability, as well as those who work in activities that are most exposed to layoffs and wage reductions and, in general, in precarious working conditions.
- Latin America is one of the most vulnerable regions to this crisis, with poverty levels of about 30 per cent and most of the population is immersed in the informal sector of the economy without social Security or safety networks. Underdeveloped or developing countries are also quite vulnerable.
- In line with increasing poverty and extreme poverty, inequality will also increase in all countries in the region. ECLAC projects increases in the Gini index from 0.5% to 6.0%.
- It is estimated that in Europe GDP growth forecasts will be negative for the euro-zone as a whole of -4% in 2020 and a positive recovery with growth of 2% in 2021.

Countries have taken other measures that mitigate and reduce the expenditure of individuals and families. As of April 24, 2020, 22 countries have announced 45 additional measures, of which 60% are credit and mortgage payment facilities, 22% are tax relief and 18% are price controls

This context led to research on the scientific and informative production of Coronavirus, COVID-19. For this, the platform used was Wizdom.ai and text mining techniques were applied to 30 articles in national newspapers in Colombia.

2. In relation to the scientific information

Figure 1 (solar ray) shows the breadth of the keyword research focus. The figure presents the different levels of study by grouping themes and sub-themes. The area where more information has been published is Biology, followed by Molecular Biology and Virology, Health Sciences and Immunology.

Figure 2 shows collaborations for this topic around the world, the darker the shade of blue, the greater the focus of research on it. It can be noted that the countries with the highest output are: The United States with 309 publications and 493 researchers in this subject, followed by Spain with 203 publications and a total of 288 researchers, in third place, is China with a total of 136 publications and 423 researchers. In Latin America, the leading country in production is Brazil with 107 publications and 329 researchers, followed by Argentina with 20 publications and 54 researchers. Reviewing the Colombian case, there are 11 publications and 15 researchers on this topic
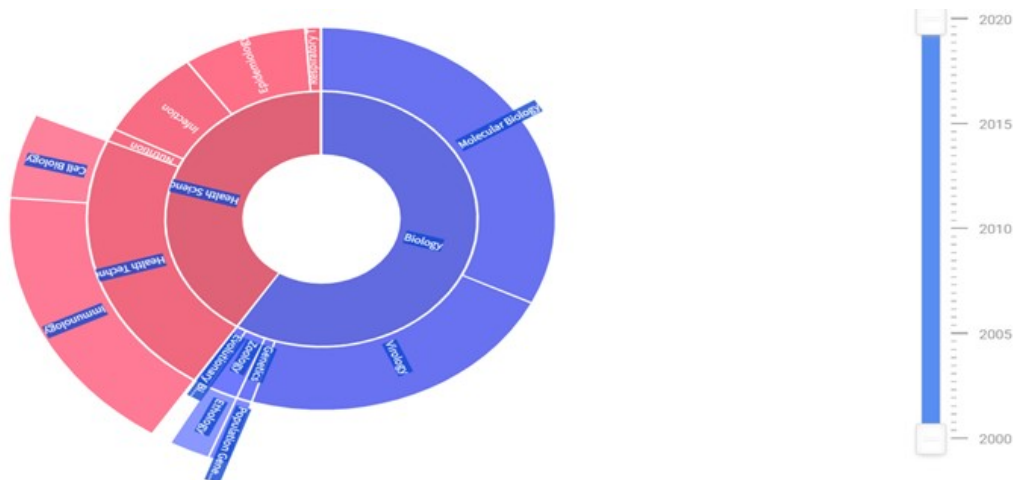
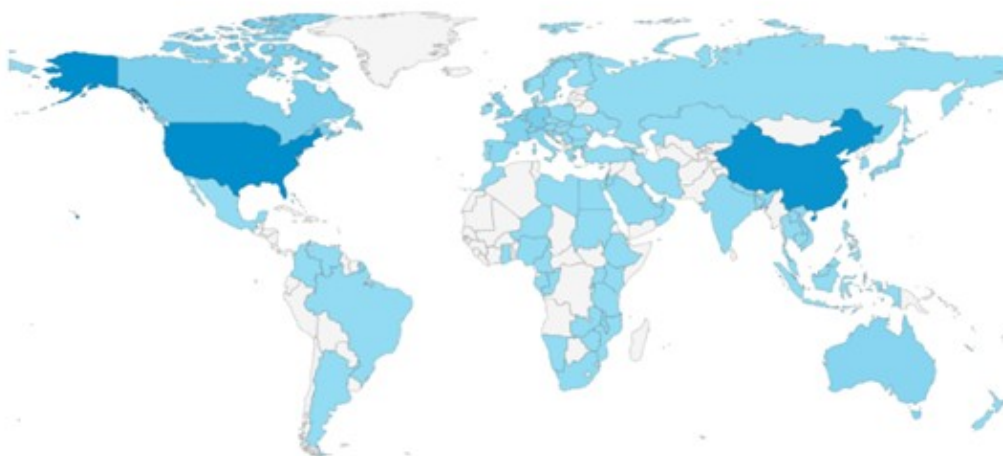Fig. 1. Production Areas. Source: Wizdow.ai

Fig. 2. Academic production. Source: Wizdow.ai

Figure 3 shows the result of publishing this topic over time. As can be seen, the term Coronavirus first appeared in 1907, from this year on, the publications on this topic are constant. In 2004 there was a growth in production when 395 documents were published, in 2010 production fell and only 196 documents were published, in 2015 a growth phase began, and 395 publications were reported in this year. By 2018, there were 364; in 2019, 421 publications and by May 2020, 425 papers had been published.
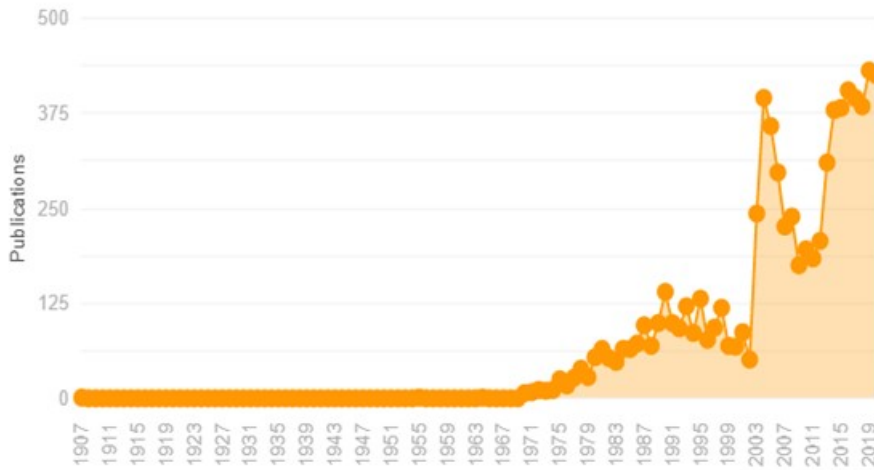
Fig. 3. Evolution of production (1907-2019) according to Wizdom.ai. Source: Wizdow.ai [3]

The word cloud (Figure 4) indicates the most frequent keywords in research for this topic. The color of the words represents different research disciplines. Figure 5 summarizes the information from the journals with the largest number of articles published on the subject. The Journal that identifies Wizdom as the most important is the Journal of Virology, ranked in quartile 1 of Scopus, and is in the American Society for Microbiology. It is followed by Virology, classified in quartile 2 of Scopus, and is from Elsevier publishing house.
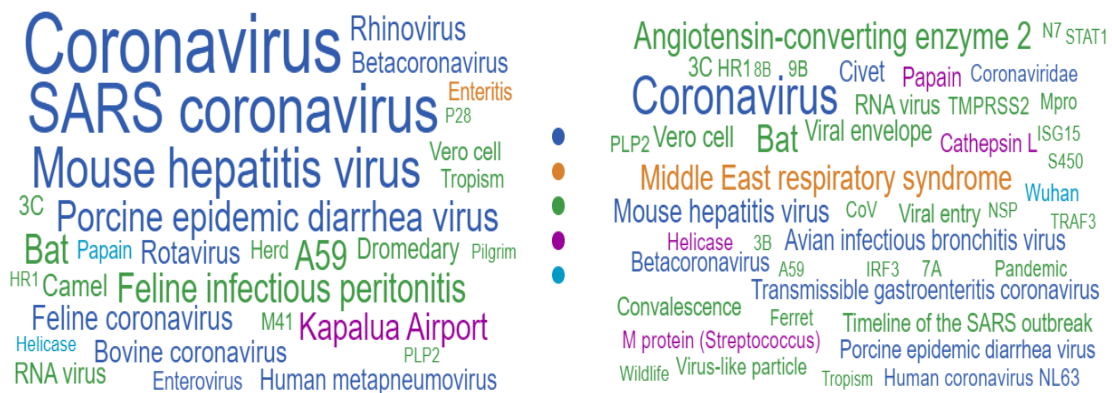

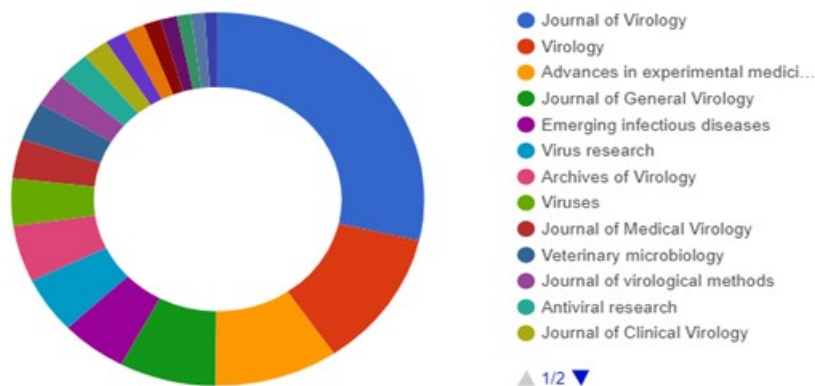Fig. 4. Topics investigated. Source: Wizdow.ai

Fig. 5. Journals with the highest number of publications, according to Wizdom.ai. Source: Wizdow.ai

Figure 6 shows the main institutions for this topic according to their total research output. First, Chinese Academy of Sciences is located with 257 publications and a share of 9.5% of the total. Second, there is University of Hong Kong with a production of 244 publications and a share of 8.5%. Third, Utrecht University produces 187 publications and has a 6.5% share. Next is the National Institutes of Health with 175 publications and a 6.1% share, followed by the University of California with 163 publications and a 5.7% share.
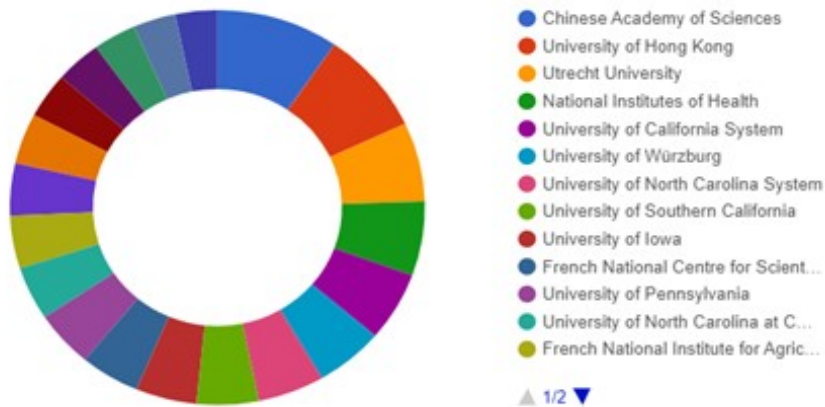

Fig. 6. Institutions leading this issue. Source: Wizdow.ai

Figure 7 shows the main authors of this topic according to their total research output. The order corresponds to Kwok-Yung Yuen with 122 publications and participation 8.4%, Luis Enjuanes with a production of 105 and a share

of the 7,2%, Ralph S Baric, with a production of 100 and a publication of 6.9%, Stanley Perlman has 91 publications and a share of 6.3%, M M with 84 publications and participation of 5.3% and Patrick C y Woo with 80 publications.
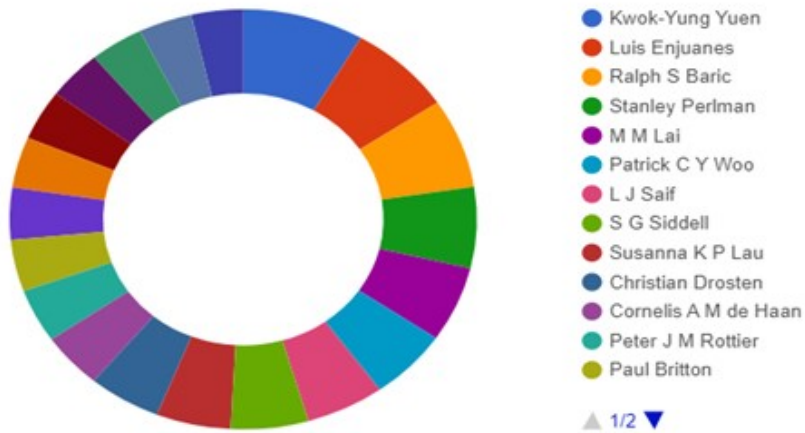

Fig. 7. Most relevant authors in this topic. Source: Wizdow.ai

The map in Figure 8 shows the geographical breakdown of patents, applicant organizations and inventors related to the subject. The country that leads this issue of patents in the subject is the United States with 609 patents, 234 organizations and 416 inventors. Second, China has 286 patents, 146 organizations and 304 inventors. Thirdly, Russia has 70 patents, 26 organizations and 137 inventors, and France is very close with 67 patents, 30 organizations and 51 inventors. Subsequently, Spain has 60 patents, 11 organizations and 26 inventors. In Latin America, Brazil has 4 patents, 6 organizations and 3 inventors, followed by Argentina with 4 patents and 2 organizations.
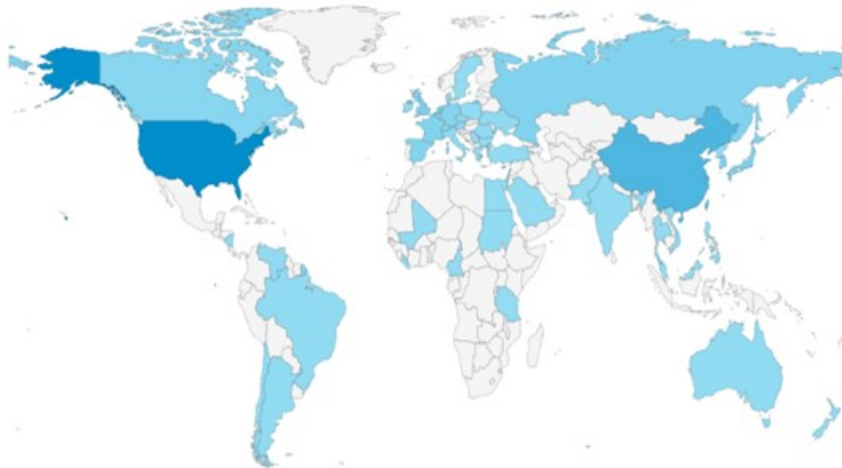

Fig. 8. Patents around this topic. Source: Wizdow.ai

3. Methodology

A qualitative analysis was carried out based on Grounded Theory, which proposes the conceptual ordering through codes derived from the data, to explain a social phenomenon [4]. Therefore, the search process considered the terms COVID - 19 and coronavirus, of which 15 articles from the source Portafolio and 15 papers from the source La República (Colombian newspapers) were chosen. The findings are conceptually based on the post-crisis management model proposed by [5] where the need for adaptation to constant change is raised.

4. Results

First, to explore the texts, a word frequency search was performed, the most frequent being Covid (0.46%), companies (0.37%), measures (0.35%), Colombia (0.30%), services (0.29%) and economy (0.26%). After cleaning the word base, a conglomerate analysis was performed with the 25 most frequent words organized into 10 groups, as can be seen in the dendrogram (Figure 9). These groups were considered for the conceptual grouping of the codes that emerged when reading line by line and will be explained below.

Fig. 9. Conglomerate analysis of the 25 most frequent words in journal papers. Source: own elaboration.

In terms of the codes associated with the measures taken from the companies for the Covid-19 is more likely the donation from the companies to the community understood as "the actions carried out that seek to strengthen different fronts: health, food, and the related value chain", as well as the work articulated between community, company and State defined as "the assistance given directly by the companies with the collaboration of the political authorities, health, family and the local police". In this way, the companies with greater market coverage made contributions to the community delivering products related to their value chain, but also contributed to the health field with the delivery of medical supplies and caps to the community, and in specific cases, inputs to the health sector, in addition to advertising communication of positive and informative messages.

However, the supplier business support code shows that large companies considered their customers to whom they supply and had to close their businesses, presenting aid bonds and various strategies to support the field. However, taking into account that the activation will be progressive, they asked the government for aid related to both

governmental and financial decisions, expressing that there will be low economic activation in the short term so they ask understanding in the payment of debts. Although they highlight the collaboration with the tax calendar change, they recognize that they are short-term measures and reinforced by the verbatim that "although they are valuable for the recovery, the time the industry sector will take to recover the levels of demand that was before will be long".

On the other hand, codes were found related to the dynamics of adaptation to the digital work that employees have had to do. Initially, there are measures that seek to preserve employees taken by employers such as the consensus of wage reduction, the implementation of telework, shift work, early holidays or others such as the modification of contracts or the reduction of personnel. Although there are also codes as to accelerate the learning process for the improvement and adaptation of products/services according to the needs understood by the rapid adaptation to the economic emergency caused by the pandemic, as well as that "the companies transform and reinvent to deliver new services that generate added value to its users". Well, this dynamic has shown committed collaborators, as well as horizontal organization understanding the company as a "permanent adaptation system managed by processes" by monitoring indicators or continuous evaluation of performance, being evident in the following intervention.

"From the responses of entrepreneurs, two elements stand out: the first one is that companies are requiring credit to maintain their operations, rather than leaning towards closing. The other one implies that they are equipping their employees, investing in technology, in rental adaptations and, above all, in changing their business models, to a large extent, because this reality that is being experienced in the world is demanding and will continue to do so in the future".

5. Conclusions

During the response to an unprecedented crisis, the Government proposes policy initiatives to protect jobs and needed economic activity, taking as a response in the national context the donation on the part of large companies leading to the joint work between Government, business and community ([5]. In this way, products/services related to its value chain were made available, as well as advertising messages and health products in response to the crisis.

[5] describes that companies performing a continuous evaluation process can "discover potential opportunities and know which improvement actions are most appropriate to make them happen" (P. 8) where, by accelerating the learning process, they show timely adaptation and immediate improvement. In the present study, adaptation to change was found relevant, from the flexibility and collaboration of people for a digital work as well as the production or distribution of products related to the crisis.

Finally, and in agreement with what was found by [6]y [7], some entrepreneurs perceive the opportunity to address current problems through business actions using available resources to create solutions to new problems or to activate available resources such as flexible payment options, joint sales initiatives, flexible staff turnover, among others; i.e. to work on identifying opportunities that create value to address the consequences of the crisis.

For future research, it would be worthwhile to have a larger documentary collection and apply other text mining techniques, such as readability indexes and sentiment analysis.

References

[1] Cifuentes-Faura, J. (2020). Coronavirus crisis: impact and economic measures in Europe and the world. Espaço e Economia. Brazilian magazine of economic geography, (18).

[2] Baldwin, R. E., & di Mauro, W. (2020). Mitigating the COVID economic crisis (No. BOOK). Centre for Economic Policy Research.

[3] PLC, I. (2019). wizdom.ai. Retrieved 12 23, 2019, from https://www.wizdom.ai/

[4] Strauss, A., & Corbin, J. (2003). L'analyse de données selon la grounded theory. Procédures de codage et critères d'évaluation. L'enquête de terrain, 363-379.

[5] Castro-Coma, M., & Martí-Costa, M. (2016). Urban Commons: from collective management to the right to the city. EURE (Santiago), 42 (125), 131-153.

[6] Kuckertz, A., Brändle, L., Gaudig, A., Hinderer, S., Reyes, C. A. M., Prochotta, A., ... & Berger, E. S. (2020). Startups in times of crisis–A rapid response to the COVID-19 pandemic. Journal of Business Venturing Insights, e00169.

[7] Torralba Quitian, O. I., & Lis-Gutiérrez, J. P. (2020). Supervised and unsupervised learning applied to crowdfunding.