

Documentary repository on the disease COVID-19

Sebastián Alfonso Gómez Lozano¹, Jorge Aurelio Herrera Cuartas²
Universidad Jorge Tadeo Lozano, Carrera 4 #22-61 Bogotá, Colombia

Corresponding author.

E-mail addresses: sebastian.gomezl@utadeo.edu.co, (S. Gómez Lozano),
jorgea.herrerac@utadeo.edu.co, (J. Herrera Cuartas) this is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

¹Universidad Jorge Tadeo Lozano, Libraries System, Bogotá, Colombia, Author.

²Universidad Jorge Tadeo Lozano, Dean of the Faculty of Economic and Administrative Sciences, Bogotá, Colombia, Thesis director.

Keywords:

Repository, Information management, Open access, COVID-19.

Abstract: The Documentary Repository on the COVID-19 disease <https://expeditiorepositorio.utadeo.edu.co/handle/20.500.12010/10356> is a bibliographic proposal that seeks to consolidate in an orderly way the scientific, normative and journalistic production on topics related to the SARS-CoV-2 virus to facilitate consultation, usability and use at different academic levels, arrange the contents, differentiating them in access with restrictions and open access in a digital repository.

1. Introduction

Digital resources in the current era are elements of easy accessibility, diffusion and dissemination. When the digital resources produced have academic or scientific relevance, it is pertinent to organize, catalog, describe, and preserve them with appropriate technologies to facilitate their retrieval and maximize usability, increasing the chances of positively impacting the search for new knowledge. Karvounarakis, G. and Kapidakis, S. (2000) stated that Digital library users can range from applications to common web browsers; these can be used to send operations to digital libraries, such as querying or browsing. Currently faced with a high proliferation of documentation on the web on topics related to COVID-19 disease, it is essential to adopt processes that allow identifying relevant documents and retrieving them promptly and accurately in the same way, prioritizing exclusively to process information and documents that are produced by reliable sources. The process of digital document management involves using a standard applied mainly by university institutions worldwide, such as a document repository, which according to the Foster Consortium (2020) "is defined as the infrastructure and service that enable the persistent, efficient and sustainable storage of digital objects (such as documents, data and codes)".

Some initiatives provide access to different contents harvested on the disease COVID-19 as the case of the Repository of the Spanish Society of Infectious Diseases and Clinical Microbiology, is a website that provides an ordered list of documents developed on the pandemic, which identifies the titles and summaries of the contents hyperlinking the objects with the original producer of the document. This resource works as a referential provider of relevant content, and it is not built on a metadata structure with a defined standard.

In turn, the Instituto Dalle Molle di Studi sull'intelligenza artificiale (IDSIA), in alliance with the Centro de Ciencias Genómicas and the UNAM have an information retrieval resource with great potential that focuses on offering services to health professionals, decision-makers, biomedical researchers, translators and IT managers, allowing access to scientific literature filtering by speciality, type, topic and subtopic. Its main contribution is to provide full-text documents from 28 medical specialities. As stated in its initial platform, this resource uses artificial intelligence procedures to consolidate the information produced. The processes include machine learning to filter and assist in the classification of the documentation.

On the other hand, the Digital Institutional Repository of the Colombian Ministry of Health partially offers information generated by the different Colombian governmental agencies, facilitating access to full-text documentation, as well as retrieval by search fields and document type; the most significant amount of information from this resource corresponds to the results of administrative acts, awards, decrees, laws and the like.

In the museographic panorama, the Ibermuseos organization has consolidated a documentary repository concerning information about the pandemic's management in museographic aspects, basically focused on reference material to disseminate the impact of the health crisis in this type of institutions. The consultation can be made in the thematic areas of impact reports on the museum ecosystem, protocols for opening museums after their precautionary closure, governmental support measures for the sector, related courses and talks.

The Organization of American States (OAS) has put into service the website Post COVID-19: Portal for Consultations, Forums and Repository; basically, the information is consolidated in two main aspects, on the one hand, there are the policies which is a space destined to host the pandemic management information created by the member countries, and on the other hand, studies which are related to the reports, analysis and reports published by scientific or academic institutions that can tend to show results and data to minimize the adverse effects caused by the Coronavirus. The operation of this resource is multidisciplinary, and documents can be accessed in different languages. It is a robust resource with thousands of documents where one of its main peculiarities is that functionally there is no free search field for the user to segment the information autonomously. In turn, it has a classification by tags where the documents are grouped thematically, which can be accessed directly to the full text or in a referential way to the portal producer of the content.

This article presents a prototype of a functional digital repository that harvests relevant and filtered information on Covid-19 from governmental and non-governmental information sources in order to consolidate it and make it available in an orderly manner in open access, to facilitate its use and retrieval, mainly to support research and academic review processes by researchers and members of academic institutions.

The results obtained are the conformation of a community of 3700 documents, grouped in 10 categories, to which it is possible to review the metrics of consultation concerning time, at the same time, it is possible to evidence the traceability of the geographic location of the IP from where the consultation is made.

Statistics show that the documents with the most significant impact are related to the efficacy percentage of the vaccines developed by the laboratories Pfizer, BioNTech, Oxford/AstraZeneca and the Russian vaccine Sputnik V against Covid-19; as well as studies

on the psychological impact caused by the child population confinement. In this same way, many consultations on the Sars-CoV-2 virus reinfection and research related to the Coronavirus similarities with other viruses such as HIV.

The document is initially organized by introducing the management of repositories and libraries and the importance of adequately managing digital information, then a chapter on information management is addressed. In section 3 the technological and regulatory guidelines are presented, which provide the foundations for the repository creation. In chapter 4 the problems encountered in the management of digital information and the multiplicity of information sources are shown. In chapter 5 the proposed development is socialized. At the end of this section, It can find the methodology of information tracking, which outlines the documentary recovery strategies used in the project. In section 7 it is presented the results obtained. In section 8 the conclusions of the development are consolidated, and finally, in section 9 the bibliography that supports the process of construction of this document is available.

2. Information management

Information management through alerts in sources has allowed the construction of the Documentary Repository Covid-19. A harvester structure of information related to the disease caused by Coronavirus, and that collects resources in academic, scientific, technical, administrative, legislative and news related aspects, the retrieval of information is mainly performed using automated RSS technology which according to Menéndez, B. (2018) "stands for *Really Simple Syndication*, a format that complies with the XML standard for sharing content on the web. It is used to disseminate updated information to users who have subscribed to a content feed."

2.2. Method of reviewing information

RSS integration allows to periodically program alerts that notify via e-mail about the production or availability of new information of interest to download, process, analyze and consolidate in the repository, in the same way when there are sources that do not support or offer RSS technology, a manual review of the information is performed, this process is common in governmental pages, where manual thematic searches are performed to identify new content. After obtaining the information, it is necessary to filter the documents, verifying the technical, technological and legal accessibility conditions of the information to minimize the possibility of infringing property rights on the works in question. During the research process, more than 3,000 highly relevant documents on the disease have been analyzed and arranged in reference or full text, obtaining usability metrics and downloading the arranged contents.

The repository's primary function is to operate as a free access point to research by linking users with information needs and providers of verified, high-quality content.

3. Guidelines for the use of libraries and technological context

3.1. Guidelines

It is important to note that due to the pandemic generated by the disease COVID-19 produced by the SARS-CoV-2 virus, government authorities have issued protocols that promote the physical use of library facilities with specific biosecurity measures to minimize exposure to the disease by users and staff, according to resolution 891 of 2020 of June 8, 2020 of the

(Ministry of Health and Social Protection of the Republic of Colombia, (2020) through which the biosecurity protocol for the management and control of the risk of the coronavirus Covid-19 in the operation of Libraries is adopted, This document generates a technical annex that normalizes the use of the Libraries where social distancing, capacity control, implementation of asepsis conditions for people and disinfection of bibliographic material prevail, since compliance with the resolution implies modifying, conditioning facilities, enabling quarantine and disinfection areas, as well as guaranteeing personal protection equipment such as gloves, gowns, masks and face masks among others in a permanent and lasting manner for collaborators and users of the information unit for the duration of the epidemiological contingency.

3.2. Technological context

From the perspective of Li, Jiao, Zhang, & Xu, (2019), "the development of digital libraries is directly related to the development of information technologies, particularly Internet technologies". According to Pérez Alarcón (2010), a digital library collects organized digital objects, which serves a defined community of users with the copyright present and managed and has preservation and preservation mechanisms. These changes are possible thanks to information technologies, according to Belloch, (n.d.) Information and Communication Technologies (ICT) have been gradually integrated in educational centers".

3.3 Use of information and communication technology

The integration of information technologies is intended to provide multimedia information, interactivity, interconnection, immateriality, more significant influence on processes, instantaneousness, digitization, penetration to various sectors and innovation. From the perspective of the information-based civilization, information technology has allowed libraries to pass through the traditional stages of automation and development of digital libraries. Thus, the systematization in the provision of library services are generated according to the technological developments that are presented, currently and more than ever in history there is an increase in the production of digital information. In turn for Colombia, according to the (DANE - National Administrative Department of Statistics, 2017) between the period of 2011-2018 went from having 23.4% to 52.7% of households with Internet connection, during this period manifests the MinTiC - Ministry of Information Technologies and Communications of Colombia, (2020) reached 10 million users with Internet connection type 4G in the country. An accelerated growth starting that in 2012 there were no users of this technology, in the same way for the year 2020, households of stratum 6 reached data download speeds on their devices of 54.0 Mbps, stratum 5 of 50.7Mbps, stratum 4 of 39.1 Mbps, stratum 3 of 29.6 Mbps, stratum 2 of 14.8 Mbps and stratum 1 of 11.0 Mbps. Consequently, these conditions increasingly facilitate access to online documentation.

3.4. Use of artificial intelligence in libraries

However, worldwide artificial intelligence is a significant technological development that allows innovations in the provision of library services, states IFLA - The International Federation of Library Associations and Institutions, (2017). Artificial intelligence will allow the following web browsers to go far beyond the analysis of keywords and evaluate websites' content. Given these circumstances, it is possible to facilitate strategies oriented to the non-presential provision of services for libraries aimed at trends in other industries and current economic models.

3.5. Digitization of information

According to ECLAC - Economic Commission for Latin America and the Caribbean, United Nations, (2018), the digitization of content, the development of transport platforms based on the IP protocol, the spread of broadband Internet and the availability of multifunctional devices have not only generated changes in the patterns of use of telecommunications services, content and information media. The volumes of information in libraries are cumulative. Usually, significant accumulations of documents are handled both physical and digital, where primarily specialized document processing and organization tools are required. From the point of view of Chen, et al. (2016), "Big data" has been a multifaceted and evolving term. It is believed that library catalog transaction records are essential sources of Big Data because they handle considerable sizes. Traditional technology has complications in supplying the volume of data, and a considerable data velocity can be evidenced, and the potential for knowledge extraction has promising viability.

Combining technologies such as Big Data, broadband, 4G internet, artificial intelligence, and cloud technologies with the information needs to be posed by the COVID-19 contingency. Suppose we add the restrictive characteristics posed in the protocols of access to libraries by the national government and institutions. In that case, we have a scenario that allows solutions to the information needs through digital media, opening possibilities for digital transformation in libraries.

4. Problem Statement

The information units require changes in their operations to offer services that minimize presentiality and guarantee access to documentation of the highest academic quality without involving high costs in acquiring bibliographic resources to be offered. This aspect is perceived more remarkably due to the impossibility of developing presential activities due to the measures adopted by the different governmental entities and to the consequences that the effects of COVID-19 may cause in society. On the other hand, different conditions tend to maximize resources, facilitate access to information and contribute to the orderly retrieval of documents related to the pandemic.

In order to confront situations of misinformation about the pandemic, society requires reliable and validated sources of information that help to minimize the possibilities of false documents and misleading news to spread, affecting decision making in the communities; these conditions are favoured due to the multiplicity of information and communication media that are within reach nowadays. Never in history has data been so accessible with the development of Web 3.0, which, as defined by *defincion.de* (2020), is related to what is known as the semantic web. Users and computers, in this framework, can interact with the network using natural language, interpreted by the software". In other words, all information hosted on the web 3.0 will be "understood" by machines.

4.1 Problems encountered

Considering the approach's argument, not only does the content created by authorities in specific disciplines come to light, but on the contrary, the content available on any website is visible and easily retrievable. This situation contributes to increasing the problem in different ways:

- Complexity in finding reliable information on the web.
- Document retrieval times can typically be longer.
- The validation of sources is a costly task that is generally performed on specific segments or selected areas of knowledge, which leads to biasing the results obtained.
- This position can be interpreted as a condition that facilitates the migration of files, changes in storage locations, and URLs' loss. In this sense, uncontrolled files or files with no traceability that guarantees the permanence of documents in a defined and permanent location make it impossible for documents to be consulted in the long term.

4.2. Opportunity to use filtered and organized information

In higher education research processes, there are different lines of exploration about COVID-19, researchers belonging to such lines should seek to find academic input expeditiously minimizing consultation times and reviewing broad spectrums of documentation, the filtering of relevant information facilitates the generation of new knowledge, this search lies in finding relevant data to continue with research and deepen on the learning of the virus, and its effects on multiple areas of knowledge, the consolidated documentary relationship in a repository on Coronavirus implies having divergent areas of knowledge converging in one place. Consequently, an appropriate implication to minimize economic costs is to generate research procedures avoiding reprocessing or replication of activities that give the same effects with different resources. It is pertinent to mention that the diligence with which information is accessed implies logistical savings for academic institutions because they concentrate their operative force on producing new results rather than on gathering documents.

4.3 Tool for the organization of information

The development of a documentary repository on COVID-19 implies implementing specialized software in the execution of repositories, managing qualified human resources to perform pertinent filtering of the acquired information, and implementing technological aids for automatic notification on the production of information on the disease that can be harvested. It is aligned with international standards of handling, description, cataloguing and document management. The developed process is adjusted to conditions of use of documents in general terms by any individual in society, whether natural or legal, without access requirements more significant than those of a stable internet connection.

4.4 Method of handling documents according to their legal aspects

The legal documents conditions imply a careful review to avoid infringing copyrights on the works exposed in the repository. Thus being respectful of the legal integrity of the documents, the possibilities and scopes that can be applied to the documents must be verified. As a general framework, the documents to be treated must have some licensing that allows the reproduction and publication; this type of documents must be of public access to impact a more significant number of users. If it is not possible to directly access the documentation, the primary source that contains and provides direct access to the document's content must be related; this process is determined in a referential way to impact the accessibility of the information directly.

5. Proposed development

The initiative to consolidate an IR arose as a strategy to support the research management

mainly of academics who research the Coronavirus that causes the disease Covid-19, which, as expressed by Bansal M., (2020), is a single-stranded RNA virus of the genus beta-coronavirus family Coronaviridae. Seven species of this family can cause infections in humans, of which four cause mainly mild respiratory symptoms and three can trigger a potentially fatal disease. The work in the IR consisted of retrieving the documentation produced by different sources worldwide. This procedure involved processing more than 2,000 documents and records over 3 months, which increased progressively.

The main page of the development includes the web embedding of the application for monitoring the development of the Johns Hopkins University pandemic, identified as COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. In the same way, the documentary integration comes from the curatorial process carried out by the World Economic Forum in the Strategic Intelligence division to monitor Covid-19 in the world and its impact on Colombia. Additionally, the relationship with other topics such as the impact on financial markets, travel, foreign trade, the spread of the virus, labour force, vaccine development, recovery from infection, among other relevant issues at the global level, is reviewed.

Image 1 shows the main interface of the Covid-19 Document Repository, with the embedded code from the Johns Hopkins Institute showing the evolution of the virus graphically worldwide.

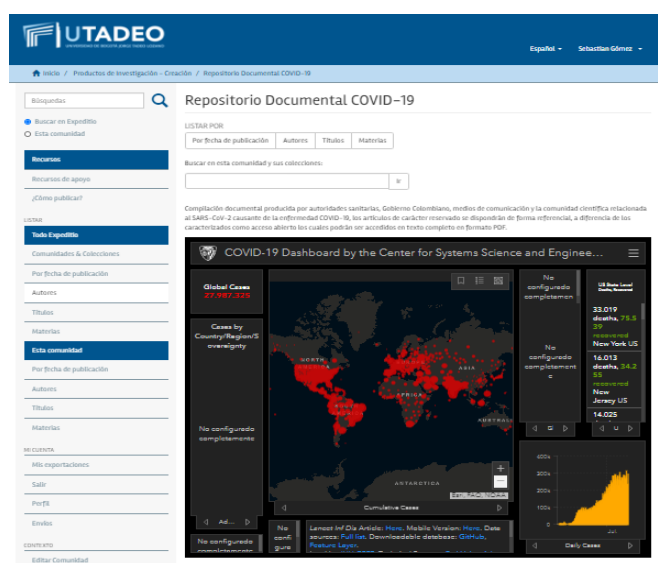


Image 1

The main interface of the IR on COVID-19

Note: The image shows the map of the evolution of the virus. Source: Johns Hopkins University & Medicine

The information harvested by the COVID-19 Documentary Repository (RI) concentrates the documents in a primary distribution in 10 collections:

- 1- Scientific papers related to COVID-19
- 2- Standards developed by the National Government
- 3- Vaccines in process
- 4- News in the media about COVID-19
- 5- Data sets on COVID-19
- 6- Books

- 7- Courses and training
- 8- Resources produced by UNESCO
- 9- Multimedia material
- 10- Mayor's Office of Bogotá

Documents produced and made available under licenses that allow preservation, bibliographic analysis and distribution in open access mode, will be shared with the same normative conditions and accessibility characteristics.

The documents with access restrictions and confidential visualization will be arranged in a referential form so that people and machines can quickly consult the direct source where the document is kept or made public.

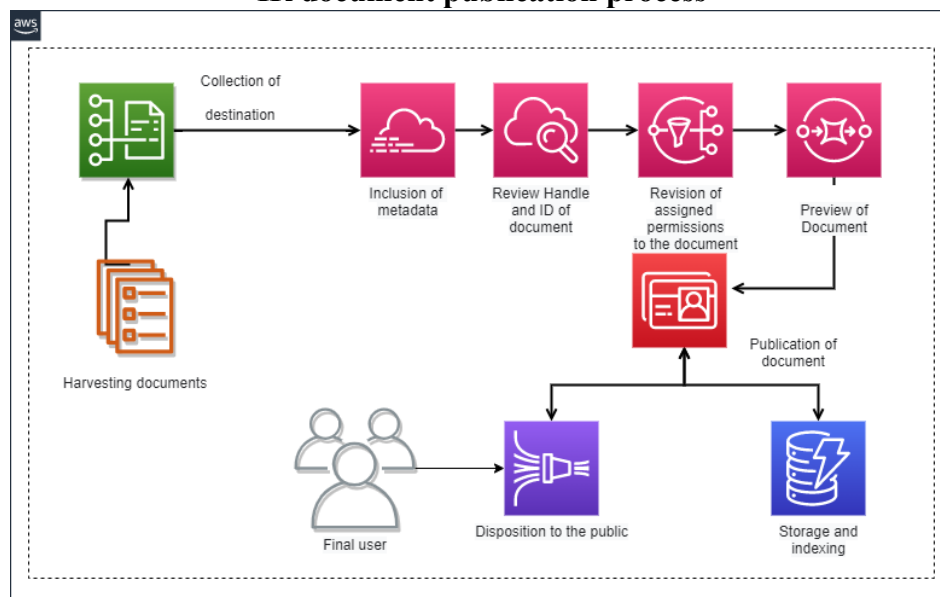
The process begins with identifying the different sources of publication of the material to be intervened. In this case, different sources of information are tracked in parallel. To select these platforms, criteria of publication periodicity, accessibility of the information, recognition and peer review of the contents are determined. Within the analysis, it is relevant to compare scientific relevance and academic and journalistic rigour, to know the list of selected information sources, see Annex 2 "Table of selected information sources to integrate documentation to the Covid-19 Documentary Repository".

The bibliographic analysis of the harvested documentation is performed using the specialized Dublin Core standard. According to Zurita, (2007), "is a simple metadata standard that attempts to describe a wide range of Web resources. The Dublin Core standard has two levels: simple and qualified. It is a set of metadata that consists of fifteen elements, which are optional and repeatable" (p.21).

Zurita (2007) defined "metadata take the form of tags or markers that help identify, describe, classify and locate all kinds of information, generally in the form of documents". (p.5). Each of the documents or elements integrated into the IR is subject to bibliographic analysis. The 15 essential metadata fields of the Dublin Core standard are identified, and these are composed in the content set by title, subject, description, source, language, relation, coverage. The second set concentrates on managing the intellectual property aspects with the fields: creator, publisher, collaborators and access rights to the document or resource being described. A third set manages the aspects of creation and identity where the metadata of date, type and format are filled. It is intended that the analyzed contents be made available in open access, a methodology that allows and facilitates access to information resources freely and free of charge. According to Sanchez, (2007), "the open access movement tries to eliminate economic and permission barriers related to access to information". (p.8). According to González, (2018), "thematic repositories collect scientific or academic documents from one or several specific scientific disciplines, and it is the researchers from various institutions who contribute by self-archiving their work." (p. 2). Image 2 shows the diagram of the publication process of the selected documents in the Covid-19 Documentary Repository.

Integration of documentation to the IR

Image 2.
IR document publication process



Note: Sequence of publication of documents from harvesting to public release for consultation.

In the development of document integration activities to the IR, the Dublin Core standard's metadata fields are analyzed to visualize those used in the development, see Annex 1.

The integration of documents into the IR is done in D-Space software, which duraspace.org states (2020) "is the software of choice for academic, non-profit, and commercial organizations creating open digital repositories. It is free and easy to install "out of the box". DSpace preserves and enables simplified, open access to all digital content types, including text, images, moving images, MPEGS, and datasets. It has a constantly growing community of developers committed to the expansion and continuous improvement of the software. Additionally, it has become one of the specific standardized systems, as stated by Morcillo López, (2016), where he indicates that "in Spain 85.71% of institutional repositories use D-Space as structural software" (p.7).

To consult the instances that allow the internal management of documents in D-Space, the consultation authorizations and permissions assigned to each document, and the resource's preview block, see Annex 3.

6. Methodology for tracking information

To achieve an automation process in retrieving information from the identified sources, Google search engine alerts (Google Alerts) were created, which (Xataka, 2020) refers to a personalized alert service based on the contents indexed and analyzed by the search engine, with the following characteristics.

- Frequency (Once a week)

- Source (finance)
- Language (Spanish)
- Region (Colombia)
- Quantity (Best results)

In the Science Direct database, a weekly alert was created with the thematic search COVID-19 with redirection of results via e-mail; additionally, in order to monitor possible vaccines for the disease COVID-19, other additional alerts have been created with the technical name of the compounds that were also monitored every week, the names of the alerts are: *INO-4800, mRNA-1273, NVX-CoV2373 and ChAdOx1 nCov-19*.

With the Elsevier database, a monitoring intervention was performed in the *Coronavirus Research Hub supports Data Scientist* to identify datasets that may be useful in the research process, to be made available on time by sharing the information in the IR. The search equation used is ("COVID-19" OR "2019-nCoV" OR "COVID19" OR "nCoV-19" OR "Sars-CoV-2").

In the Nature Research platform, advanced searches and monthly information filtering were performed to harvest research resources of thematic type by COVID-19, type of information (Research), temporal anticipation of (Last 30 days) considering the products Nature Communications, Scientific Reports, British Dental Journal, Nature and Signal Transduction and Targeted Therapy.

They are using the data.mendeley.com platform and filtering by *Sequencing Data* to obtain the corresponding information and by Dataset to retrieve the data compilations to download and filter information to transfer and redirect from the IR to the different data compendia. In the Directory of Open Books source, manual and periodic reviews are performed to validate new provisions of book-type materials that are consolidated as prospects to be integrated into the IR; in the same way, manual crawls are applied in the UNESCO platform to review new provisions and documentary productions related to COVID-19 that may be subject to retrieval by the IR.

7. Results obtained

As of March 11, 2021, 3719 documents have been analyzed, described, catalogued and entered into the COVID-19 Document Repository, with the following distribution:

- Multimedia material: 1
- Courses and training: 1
- Books: 37
- UNESCO: 11
- COVID-19 Datasets: 2
- Scientific documents related to COVID-19: 2.292
- Standards developed by the Colombian National Government: 364
- Vaccine documents in process: 73
- News in the media about COVID-19: 894
- Office of the Mayor of Bogotá DC 90 Queries

in the application yield the following metrics:

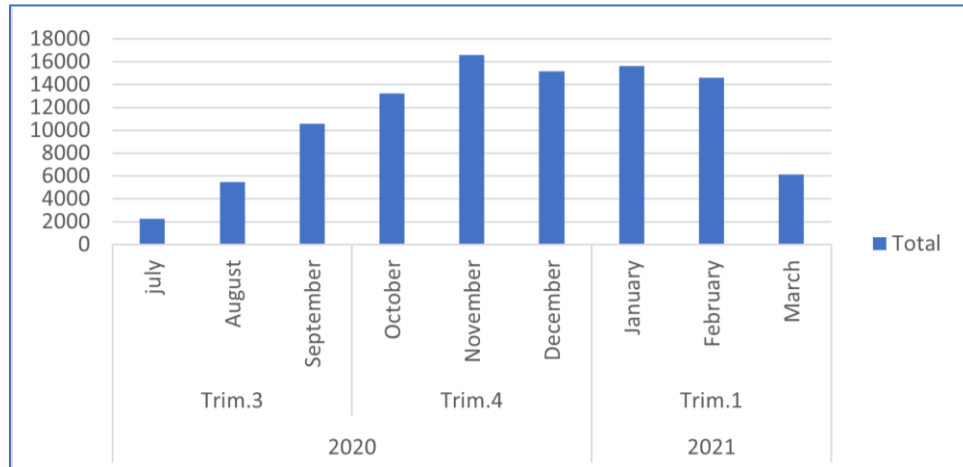
From July 9, 2020, to March 11, 2020: 101,753 queries, 79,632 were geographically located,

resulting in a daily average of 370 queries in 275 days of the project.

Image 3 shows the evolution of queries in the Covid-19 Document Repository from the publication and public access process's initial date.

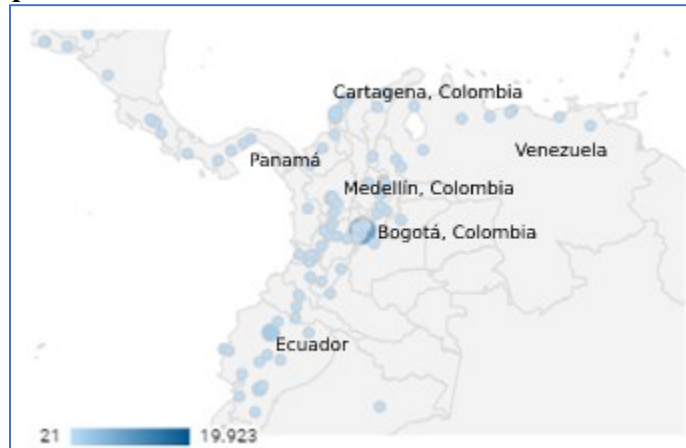
Image 3.
RI consultations per day

Note: Consultation figures of the documentation hosted in the repository.



From the initial date of the documents' availability to the public until March 11, 2021, we identified 2. The cities identified within the top 20 with the most queries are composed of Bogota, Mexico City, Quito, Santiago, Medellin, Nuestra Señora de la Paz, Madrid, Buenos Aires, Cuenca, Soacha, Arequipa, Barranquilla, Merida, Guatemala City, Barcelona, Santa Cruz de la Sierra, Mexicali, Arjona, Cali and Turbaco, accounting for 38% of the total number of queries. Bogota, with 8,354 queries, is the population with the highest number of document views in the repository. Image 4 shows the geographic distribution of queries on Colombia and neighbouring territories.

Image 4.
Geographic reference of the cities from which the RI is consulted.



Note: Data of the cities with the most document queries in the repository.

Table 1 shows the corporate authors' list with the most significant number of documents in the repository.

Table 1. Corporate authors

Note: Data were taken from the repository query interface. Table 2 shows the natural authors' list with the most significant number of documents in therepository.

Source	Number of documents harvested
Periódico El Tiempo	83
Ministerio de Hacienda y Crédito Público	57
Revista Semana	47
CNN en Español	45
Ministerio de Salud y Protección Social	34
Ministerio de Defensa Nacional	32
Periódico El Espectador	31
Revista Dinero	30
Ministerio del Interior	26
Deutsche Welle	25
Diario La República	21
Emisora Caracol Radio	21
Ministerio de Comercio, Industria y Turismo	19
Ministerio de Educación Nacional	19
Periódico El Heraldo	19
BBC News	15
Instituto Colombiano de Bienestar Familiar	14
Periódico El Colombiano	14
Ministerio de Justicia y del Derecho	13
Ministerio del Trabajo	13
Periódico El País de Cali	13
Canal RCN	12

Diario La Vanguardia	12
Jurisdicción Especial para la Paz	12
Periódico Portafolio	12

Table 2. Natural authors

Note: Figures of the most relevant authors entered in the repository.

Author	Number of documents	Total citations	H index
Misra, Anoop	11	31436	86
Vaishya, Raju	8	2292	23
Liu, Yang	7	1039	18
Deng, Wei	6	988	17
Iyengar, Karthikeyan	6	196	8
Qin, Chuan	6	2458	27
Gupta, Ritesh	5	29246	75
Haleem, Abid	5	4987	34
Hu, Yi	5	131	7
Javaid, Mohd	5	1077	19
Jiang, Shibo	5	22916	80
Khan, Suliman	5	1521	14
Li, Jing	5	22	2
Liu, Shuwen	5	14471	45
Pranata, Raymond	5	297	9
Rezaei, Nima	5	25556	68
Siddique, Rabeea	5	527	10
Singh, Ritu	5	712	14

Table 3 shows a list of the keywords with the most occurrences in the repository.

Table 3.

Keywords	Number of occurrences
COVID-19	2270
Coronavirus	1896
SARS-CoV-2	1896
Síndrome respiratorio agudo grave	1890
Coronavirus – Jurisprudencia -- Colombia	360

Recurring keywords

Note: Key terms assigned to documents entered into the repository.

Table 4 provides a list of the most frequently used search terms in the repository.

Table 4.

Data were taken from the Repository query interface.



Ranking	Search terms
1	Misra, Anoop
2	Chen, Lei
3	Chen, Zhichao
4	Chowell, Gerardo
5	Drosten, Christian
6	Ministerio de Comercio, Industria y Turismo
7	Ministerio de Hacienda y Crédito Público


Table 5 identifies the Top 3 of the most consulted documents in the IR with Almetric indicators that, according to the (University Library of Huelva, 2020), are the alternative indicators based on the web 2.0. or social web and that deal with the analysis of the activity and visibility of scientific and academic production.

Table 5

Top 3 most consulted documents

Note: Data obtained from the repository query interface and each of the items identified by selecting the respective Almetrics.

Citation APA	Consulted	Almetric
Pérez, A., et al. (2020). <i>Impacto psicológico del confinamiento en la población infantil y como mitigar sus efectos: revisión rápida de la evidencia</i> . Recuperado de: http://hdl.handle.net/20.500.12010/13278 .	3099	 Twitter (21) Mendeley (3)
Ruiz, M. T. (2020). <i>Las estadísticas sanitarias y la invisibilidad por sexo y de género durante la epidemia de COVID-19</i> . Recuperado de: http://hdl.handle.net/20.500.12010/13278 .	2803	 Twitter (11)

<u>010/13274.</u>		
Tang, S., et al. (2020). <i>Aerosol transmission of SARS-CoV-2? evidence, prevention and control</i> . Recuperado de: http://hdl.handle.net/20.500.12010/11891	63	 <u>News (71)</u> <u>Blogs (3)</u> <u>Twitter (382)</u> <u>Wikipedia (1)</u> <u>Reddit (1)</u>

8. Conclusion

Considering the consultation indicators of the documentation available in the repository, it is crucial to conclude that the volume of consultations is directly proportional to the number of documents deposited since the storage process is continuous and cumulative over time. The grouped arrangement of documents by thematic axes on COVID-19 represents a significant input for knowledge management with a high potential to positively impact the creation of new academic, administrative or scientific content. Documentary treatment based on standards defined and adopted by the academic sector allows for quick and orderly dissemination of knowledge through digital media and the autonomous retrieval of information through the web by interested users. Documentary consolidation allows traceability on the usability of science by determining technical measurements on the impact at general levels of the consulted documents.

A Documentary Repository on COVID-19 is a filter capable of contributing to the proliferation of false and misleading news about the disease. Simultaneously, it has effects in providing support on the informative ordering of essential topics of interest to society. Considering the logistic implication necessary to sustain the documentation hosted and at the service of society in general, the IR works as an agent of social responsibility, being an inclusive, unrestricted and altruistic application to contribute in documentary aspects to alleviate the crisis caused by the Coronavirus.

Adequate management of documents related to COVID-19 in the IR helps to avoid reprocesses in the creators of scientific and/or administrative information, given that it is possible to evidence the most significant number of consultations in Colombian and Hispanic American cities, among them several intermediate or small cities that benefited from having the possibility of making better-documented decisions about the course of the pandemic.

9. Bibliography

Foster Open Science. (2020). Glosario. From Foster Open Science: <https://book.fosteropenscience.eu/es/06Glossary/>

Bansa, M. (2020). Cardiovascular disease and COVID-19. From PubMed: <https://pubmed.ncbi.nlm.nih.gov/32247212/>

Belloch, C. (2012). Las Tecnologías de la Información y Comunicación en el aprendizaje. Material docente [on-line]. From Departamento de Métodos de Investigación, Universidad de Valencia: <https://www.uv.es/bellochc/pedagogia/EVA1.pdf>

Biblioteca Universitaria de Huelva. (2020). Altmetrics: ¿Qué son las altmetrics? From guiasbuh.uhu.es: <https://guiasbuh.uhu.es/altmetricas>

CEPAL - Comisión Económica para América Latina y el Caribe, Naciones Unidas. (2018). La nueva revolución digital, de la internet del consumo a la internet de la producción . From Repositorio Cepal: https://repositorio.cepal.org/bitstream/handle/11362/38604/4/S1600780_es.pdf

Chen, H.- l., Doty, P., Mollman, C., Niu, X., Yu, J.- c., & Zhang, T. (2016). Library assessment and data analytics in the big data era: Practice and policies. From asis&t: <https://asistdl.onlinelibrary.wiley.com/doi/full/10.1002/pra2.2015.14505201002>

DANE - Departamento Administrativo Nacional de Estadística. (2017). Encuesta nacional de calidad de vida. From <https://www.dane.gov.co/index.php/estadisticas-por-tema/salud/calidad-de-vida-ecv/encuesta-nacional-de-calidad-de-vida-ecv-2017>

Definición.de. (s.f). WEB 3.0. From Definición de: <https://definicion.de/web-3-0/#:~:text=La%20web%20se%20vincula,se%20conoce%20como%20web%20sem%C3%A1ntica.>

Digital, R. I. (2020). MinSalud . From Ministerio de Salud : <https://www.minsalud.gov.co/sites/rid/Paginas/buscar.aspx>

Gonzalez-Argote, J. (2020). Repositorio de investigaciones estudiantiles: tarea necesaria y trascendentalRepository of student research: A necessary and important task. From ScienceDirect : <https://www.sciencedirect.com/science/article/pii/S1575181318301396>

Iber Museos . (2020). El espacio de los Museos iberoamericanos . From Ibermuseos : <http://www.iber museos.org/>

IFLA - The International Federation of Library Associations and Institutions. (2017). Advances in Artificial Intelligence. From <https://trends.ifla.org/literature-review/advances-in-artificial-intelligence>

Karvounarakis, G. & Kapidakis, S. (2000). Submission and repository management of digital libraries. From doi: 10.1016/s1389-1286(00)00157-2

Li, S., Jiao, F., Zhang, Y., & Xu, X. (Enero de 2019). Problems and Changes in Digital Libraries in the Age of Big Data From the Perspective of User Services. From science direct: <https://doi.org/10.1016/j.acalib.2018.11.012>

Menéndez-Barzanallana, A. R. (2018). Prácticas: lectores RSS, Feedly. From Universidad de Murcia : <https://www.um.es/docencia/barzana/PRACTICAS/RSS-Google-Reader.html>

Ministerio de Salud y Protección Social de la República de Colombia. (2020). Resolución Número 000891 de 2020. From MinSalud : https://www.minsalud.gov.co/Normatividad_Nuevo/Resoluci%C3%B3n%20No.%20891%20de%202020.pdf

MinTiC - Ministerio de Tecnologías de la Información y las Comunicaciones de Colombia. (2017). Plan TIC 2018 - 2022 El futuro digital es de todos. From mintic.gov.co: https://www.mintic.gov.co/portal/604/articulos-101922_Plan_TIC.pdf

Morcillo López, L. (2016). Los repositorios institucionales en las universidades públicas de España: estado de la cuestión. From Cuadernos de Gestión de Información : <https://revistas.um.es/gesinfo/article/view/264121>

OEA. (2020). OEA. From <http://www.oas.org/es/>

Pérez Alarcón, A. (16 de 02 de 2010). La biblioteca digital. From Repositorio Institucional Universitat Oberta de Catalunya: <http://openaccess.uoc.edu/webapps/o2/handle/10609/310>

Sánchez, N. (2007). El movimiento de acceso abierto a la información y las políticas nacionales e institucionales de autoarchivo. From Scielo: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352007000900005

Xataka. (2020). Google Alerts: cómo activarlo y 14 alertas para sacarle el máximo partido. From Xataka: <https://cutt.ly/XhY0LGX>

Zurita Sánchez, J. (2007). Introducción a Dublin Core. Mexico D.F: Ciudad Universitaria.