



Las opiniones y los contenidos de los trabajos publicados son responsabilidad de los autores, por tanto, no necesariamente coinciden con los de la Red Internacional de Investigadores en Competitividad.



Esta obra por la Red Internacional de Investigadores en Competitividad se encuentra bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivadas 3.0 Unported. Basada en una obra en riico.net.

Los datos abiertos pos-COVID-19: perspectiva a partir de la producción científica

Carlos Estrada Zamora¹

Resumen

Los datos abiertos también se vieron afectados por la pandemia por COVID-19 a partir del año 2020, así lo demuestra la producción científica desde esa fecha, por lo que este trabajo analiza las representaciones gráficas generadas por la herramienta VOSviewer a través de registros semánticos extraídos del repositorio Scopus, su interrelación y temporalidad para la identificación de tendencias en perspectiva en los años posteriores al Coronavirus SARS-COV-2; los metadatos apuntan al interés por la investigación problemas sociales y del entorno, vinculados a los problemas propiciados por la crisis que generó la pandemia y la integración creciente de estos con las cuestiones pre-pandemia.

Abstract

Open data were also affected by the COVID-19 pandemic from 2020, as evidenced by the scientific production since that date, so this paper analyzes the graphical representations generated by the VOSviewer tool through semantic entries extracted from the Scopus repository, their interrelation and temporality for the identification of trends in perspective in the years following the SARS-COV-2 Coronavirus; the metadata suggest an interest in research on social and environmental problems, related to the problems caused by the crisis generated by the pandemic and the growing integration of these with pre-pandemic matters.

Keywords: Datos Abiertos, Ciencia de datos, Gestión del conocimiento, Cienciometría, COVID-19

¹ Universidad de Guadalajara.

Introducción

Este trabajo, tiene como objetivo la identificación de tendencias a partir de la bibliografía de trabajos científicos contenidos en repositorios de compilación en línea de prestigio (como Scopus) en el uso de los datos abiertos durante la pandemia por COVID-19 (diciembre de 2019 al año 2021), así como las perspectivas de la producción científica de la temática para la era de transformaciones que se avecina a partir de las crisis generadas por este fenómeno.

El entonces, nuevo coronavirus (2019-nCoV / COVID-19), provocado por la acción del virus SARS-CoV-2, que provocaba una neumonía desconocida (o atípica), comenzó su camino por el mundo de acuerdo con los reportes oficiales, a partir de diciembre de 2019 a la sombra del invierno en el hemisferio norte, las primeras transmisiones entre personas se documentaron en la ciudad de Wuhan, Hubei (China), hasta alcanzar prácticamente todos los rincones de la tierra (más de 150 países) (Malhotra et al., 2019; Wang et al., 2020; World Health Organization, 2020).

El rápido progreso de la infección estremeció a todas las sociedades en el planeta entero, con efectos (positivos y negativos) en todos los ámbitos: desde la salud hasta la economía, que, sin duda, han transformado profundamente los procesos vitales contemporáneos. Entre los efectos negativos de la pandemia, la salud, con la saturación de hospitales y la falta de tratamientos médicos y vacunas fue una de las vertientes más sentidas y que movilizó a un gran número de organizaciones para la unión de recursos para su gestión. Desde luego, la economía mundial vista desde el comercio recibió un fuerte golpe debido a los cierres de negocios a causa del confinamiento y la limitación de espacios para la asistencia de personas, lo que contrajo el gasto y el consumo ante el temor de los usuarios (Shin y Vandembroucke, 2022).

Asimismo, la educación, la ciencia, la política, el turismo, el entretenimiento y otros ámbitos en la vida de las personas sufrieron afectaciones y, por supuesto, cambios que aún no han terminado por definirse y aún representan nerviosismo y preocupaciones para los involucrados (Damaševičius y Zailskaitė-Jakštė, 2022; Hamrouni et al., 2022; Song, 2022).

Por otra parte, no ha sido todo negativo, también se desarrollaron importantes cambios en términos, por ejemplo, de comercio, con la consolidación de plataformas digitales de venta y la conformación de cadenas de suministro y logística más robustas (Miljenović y Beriša, 2022), la educación también con importantes avances en digitalización y nuevos métodos de enseñanza que han hecho más incluyente al conocimiento a través de las aulas (ahora también virtuales) (Simuț et al., 2021), y un sinnúmero de avances también en medicina, desarrollo humano emocional, las condiciones de trabajo de las personas, entre otros (Pujo *et al.*, 2022).

En este contexto tan convulso por las circunstancias de la pandemia por COVID-19, donde las tecnologías de la información tuvieron un papel preponderante, los datos que permiten tomar (mejores) decisiones fueron imprescindibles. Aquí, los datos abiertos que bien definen Aleixandre-Benavent *et al.* (2021) son todos aquellos datos o información que puede ser reutilizada y distribuida sin restricciones (salvo la atribución de los mismos) para que cualquier individuo u organización pueda hacer uso de ellos para su beneficio.

Datos abiertos en pandemia

En esta primera instancia, se documenta el uso de los datos abiertos a partir de los trabajos científicos que se desarrollaron durante la pandemia por COVID-19 desde su inicio (diciembre 2019) hasta el fin de las primeras fases de confinamiento (inicios del año 2021), esto a partir de los registros en la plataforma Scopus que recopila la información sobre dicho contenido científico a partir de fuentes de prestigio que guardan trabajos expuestos a revisiones rigurosas y de gran calidad.

Para lograr este primer objetivo, vale la pena definir que la primera fase de la pandemia por COVID-19 vivida desde su descubrimiento y expansión desde diciembre de 2019, consistió en eventos de confinamiento y la implantación de medidas extremas que permitieron a los gobiernos de prácticamente todos los países del mundo, frenar la diseminación del coronavirus (como también se le conoció) entre los habitantes (Chen *et al.*, 2022). Esto, en mayor o menor medida, se mantuvo hasta la mitad del año 2021, donde gracias a la expansión de la vacunación y las estrategias de sanidad en los espacios de convivencia de las personas, fue seguro (gradualmente) retomar algunas actividades de lo cotidiano pre-pandemia (Sohel y Md. Rabiul, 2022).

Scopus, es una base de datos propiedad de Elsevier que recopila información bibliográfica sobre diversas vertientes de la ciencia (campos) y las pone a disposición en su portal (<https://www.scopus.com>) (Elsevier B.V., s/f), a partir de esta fuente, se encontraron 35 registros relativos a trabajos cuyo título contiene las palabras “COVID-19”, “SARS-COV-2” y “Open Data”, los cuales abarcan diversos tópicos que se mencionan más adelante.

Con el propósito de desvelar los usos de los datos abiertos durante la pandemia de acuerdo con los registros sobre los trabajos científicos contenidos en Scopus, se localizaron trabajos como el de McClary-Gutierrez *et al.* (2021) el cuál aborda la importancia de compartir “metainformación” sobre el análisis de aguas residuales para la detección de COVID-19, o el trabajo sobre la relación de la diseminación del virus en consonancia con el clima de Dipta y Naoki (2021), así como la investigación de Ali *et al.* (2022), sobre los formatos no homogenizados de publicación de datos abiertos por parte de organizaciones del mundo, lo cual dificultaba la reutilización de los mismos para la toma de decisiones sobre riesgos sanitarios. Estos trabajos implicaron esfuerzos desde la ciencia

para la comprensión y combate del propio virus como estrategia que aporta conocimientos sobre el riesgo para la salud que representó.

Por otra parte, el esfuerzo que documenta el artículo de Wahltinez *et al.* (2022) sobre una herramienta en la que la empresa de tecnología Google, pone a disposición de las personas un enorme repositorio de datos epidemiológicos útiles y confiables sobre COVID-19 a partir de fuentes de información oficiales y con alto rigor en cuanto a su autenticidad y confiabilidad para su reutilización. En el mismo sentido, el trabajo de Strcic *et al.* (2022) analiza la disponibilidad de información sobre COVID-19 durante la etapa más dura de la pandemia (los años 2020 y 2021) en cuanto a la apertura de los microdatos respecto a los cuales se realizaron análisis que derivaron en trabajos científicos y la importancia de que éstos, se encuentren a la disponibilidad del público (Gkiouras *et al.*, 2020; Pecoraro y Luzi, 2021).

Desde luego, los datos abiertos contribuyeron en iniciativas como las que describen Benning *et al.* (2021) sobre el transporte público como herramienta para monitorizar contagios de COVID-19, también como medios para monitorizar las consultas relacionadas a la incertidumbre sobre las escuelas y la educación en tiempos de pandemia (Eugene *et al.*, 2021), también proyectos de innovación social de los que habla Almeida (2021) en su trabajo, los cuales no sólo se centran en las capacidades que tenían los datos abiertos para ayudar a las personas en situaciones colaterales a la pandemia, si no en la diseminación de servicios que ayudaran en tareas de acceso a la información, servicios turísticos, entre otros.

Los trabajos de Kobayashi *et al.* (2021) y Hagen *et al.* (2021) se enfocaron en la necesidad que se tuvo de datos abiertos como vehículos contra la infodemia (diseminación de gran cantidad de información sin verificación de autenticidad) y para la generación de trabajo en red para la reproducción de tecnología cívica con miras a la inteligencia colectiva.

La exploración de los trabajos contenidos en las listas de Scopus que fueron expuestos en este apartado, son muestra de la amplia y emergente capacidad de los científicos de los datos abiertos para hacer frente a los retos que supuso la pandemia por COVID-19 en los años 2020 a 2022 y propiciar bien social.

Primer vistazo a los datos abiertos pos-pandemia: el análisis de la perspectiva a partir de la producción científica

Luego de un primer recorrido sobre los (35) trabajos científicos en Scopus en el ámbito de los datos abiertos y su uso durante la pandemia por COVID-19, se desarrolla para el presente estudio de revisión bibliométrica, la metodología y el análisis del mismo repositorio durante el mismo periodo

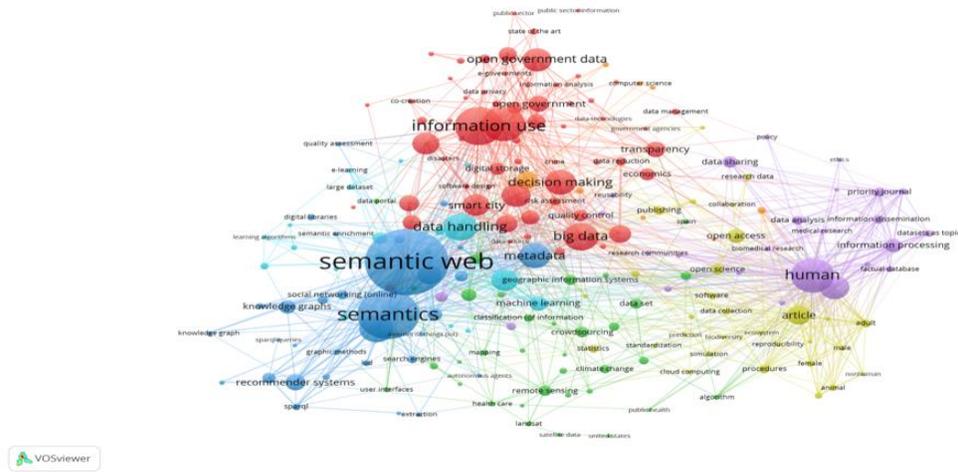
de tiempo la producción completa acerca de “open data”, con la finalidad de identificar las tendencias en los trabajos durante este tiempo y estimar posibles líneas de investigación futuras.

El análisis bibliométrico (o cienciométrico), consiste en la visualización gráfica de indicadores a partir de la co-ocurrencia semántica en un repositorio o base de datos sobre producción científica (Petrovich, 2022; van Eck y Waltman, 2014), y con todo esto, generar métricas para el análisis de esa información de acuerdo con un contexto de interés científico o por simple analítica. Para la ejecución del análisis en mención, se emplea la versión 1.6.18 de la paquetería informática de libre uso desarrollada por Centre for Science and Technology Studies (s/f) de la Universidad de Leiden (Países Bajos), VOSviewer, la cual permite generar mapas basados en datos que se interrelacionan con diferentes niveles de detalle (van Eck y Waltman, 2010).

Tomando como base el repositorio de información de Scopus, se eligieron los trabajos científicos etiquetados en el título con el término “Open Data” (sin importar, que estos aborden o no la temática del COVID-19) y que se publicaron en los años 2021 y 2022, periodo en el cual, se podrían considerar producciones posteriores a la época del confinamiento. La búsqueda arrojó 498 registros al 21 de agosto de 2022.

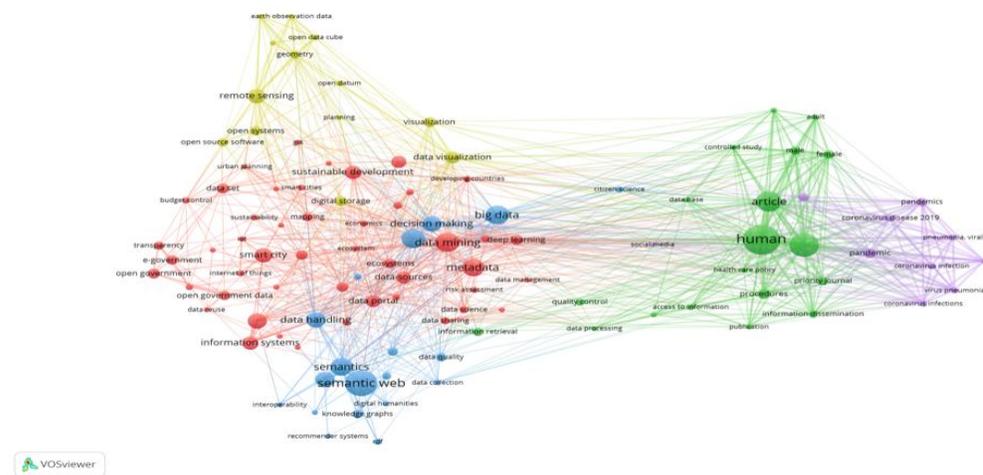
Para la realización del análisis correspondiente a este estudio, se recurrió a la ejecución de mapas de visualización de datos en red en VOSviewer, mediante los análisis de co-ocurrencia (considerando un mínimo de 5 ocurrencias) a partir de 3,886 palabras clave que se generaron en la base de datos de la producción científica de los años 2021 y 2022 que se encuentran en Scopus con la temática en la que este trabajo se enfoca. A la par, y para hacer latentes las diferencias, se ejecutaron los mismos análisis de clústeres en las bases durante la pandemia, año 2020 (con 2,989 palabras clave), y antes de la pandemia, los años 2018 y 2019 (con 4,734 palabras clave). A partir de ello, se generaron los registros gráficos que siguen (véanse de la Figura 1 a la Figura 5, consecutivamente).

Figura 1. Clústeres de redes semánticas a partir de registros bibliográficos de Scopus durante los años 2018 y 2019 con base en palabras clave con al menos cinco ocurrencias sobre el término “Open Data” al 21 de agosto de 2022. Elaboración mediante VOSviewer.



En la Figura 1, correspondiente a la producción previa a la pandemia por COVID-19 en los años 2018 y 2019, destacan trabajos científicos relacionados con la web semántica (con tópicos enfocados en el adelanto del conocimiento para la tecnología), el uso de la información para la toma de decisiones (principalmente en temas de mejora gubernamental) y el desarrollo humano (donde también existen fuertes conexiones en la producción científica en el avance de tecnologías para las personas).

Figura 2. Clústeres de redes semánticas a partir de registros bibliográficos de Scopus durante el año 2020 con base en palabras clave con al menos cinco ocurrencias sobre el término “Open Data” al 21 de agosto de 2022. Elaboración mediante VOSviewer.



Discusión

Tomando como base los resultados expuestos en la sección anterior de esta obra, se puede visualizar en primera instancia y como elemento destacable para su examinación, la disgregación de la producción científica durante la época de confinamiento e inicio de la pandemia a lo largo del año 2020 (véase Figura 2), donde los esfuerzos de divulgación científica se concentraron de manera importante en los datos abiertos como vehículo de propuestas para el conocimiento con el fenómeno latente del COVID-19.

Ya en la representación gráfica pos-COVID-19 (la de los años 2021 y 2022) (véanse figuras 3 y 4), se puede apreciar la reintegración de las relaciones de los trabajos científicos, pero con la latencia del Coronavirus integrándose en la producción y divulgación del saber. La Figura 5, hace evidente la preponderancia de trabajos sobre datos abiertos e iniciativas muy relacionadas al factor humano, donde el bienestar (y sus temáticas asociadas) se posiciona como tendencia significativa para el estudio a futuro del *open data*.

En definitiva, y acorde con la gráfica de la Figura 5, la perspectiva pos-pandemia se dirige a retomar el ritmo hacia la gestión de la información, el *machine learning* y la apertura de la ciencia, con la integración de novedosas y más amplias fuentes de datos (como estadísticas, imagen satelital, datos climáticos, entre otros). Y, aunque se retoma el camino, se identifican múltiples relaciones marcadas por el COVID-19, pero menos centradas en esta pandemia, que podría muy pronto convertirse en endemia.

Los trabajos sobre uso de imágenes satelitales para determinar los cambios en la vegetación en zonas de África (Wardle *et al.*, 2022), la deforestación en Italia (Francini *et al.*, 2022) o el uso de estadísticas para determinar deficiencias en el suministro eléctrico en Alemania frente a la proliferación de autos eléctricos (Straub *et al.*, 2021), son algunos de los ejemplos que demuestran las tendencias implicadas en el párrafo anterior.

Conclusiones

Las herramientas para la analítica de metadatos como VOSviewer permiten el re-procesamiento de compendios de información y de esta manera, conocimiento nuevo, mismo que es muy productivo porque permite echar una mirada en el cúmulo de conocimientos. Esto también surge como una advertencia acerca de la trascendencia de la apertura de los datos para generar innovación social, lo que no sólo corresponde a iniciativas desde el gobierno, si no el involucramiento de los entes privados.

La irrupción del COVID-19 en las sociedades del mundo también ha afectado el desarrollo de la ciencia, haciéndola más consciente del entorno y enfocándose aún más en los problemas sociales y las soluciones que reviertan las crisis presentes y futuras. Será tarea de los investigadores no desestimar lo aprendido ante el Coronavirus y concebir esquemas que blinden los recursos básicos ante las adversidades posibles en el porvenir.

Los próximos años en registros de productos científicos demostrarán que las investigaciones se encausaron hacia las tendencias aquí plasmadas con base en la evidencia, o se retomaron a los asuntos pre-pandemia sin los componentes que los hacen ricos ahora. El fortalecimiento y no el retroceso, será la clave para la construcción de un panorama de datos abiertos más robusto e incluyente.

En el pos-COVID-19, los datos abiertos se convierten en herramienta multisectorial capaz de fortalecer la tarea científica de prácticamente todas las áreas del conocimiento conforme el acceso a la información se transforma en gestiones para el beneficio colectivo y el enriquecimiento de los puntos de vista donde confluyen los saberes de disciplinas tradicionalmente opuestas.

Referencias

- Alexandre-Benavent, R., Ferrer Sapena, A., y Peset, F. (2021). Compartir los recursos útiles para la investigación: Datos abiertos (open data). *Educación Médica*, 22, 208–215. <https://doi.org/10.1016/j.edumed.2019.07.004>
- Ali, S., Zada, I., Mehmood, Z., Ullah, A., Ali, H., y Ullah, M. (2022). Publishing and Interlinking COVID-19 Data Using Linked Open Data Principles: Toward Effective Healthcare Planning and Decision-Making. *Mathematical Problems in Engineering*, 2022, 1–16. <https://doi.org/10.1155/2022/4792909>
- Almeida, F. (2021). Open Data's Role in Social Innovation Initiatives to Fight COVID-19. *Central European Management Journal*, 29(3). <https://doi.org/10.7206/cemj.2658-0845.51>
- Benning, O., Calles, J., Kantarci, B., y Khan, S. (2021). Transit Networks, Social Contacts, and Open Data Meet Public Transportation Plans for Post-COVID-19: A Canadian Case Study. *IEEE Engineering Management Review*, 49(3), 30–41. <https://doi.org/10.1109/EMR.2021.3103158>
- Centre for Science and Technology Studies, L. U. (s/f). *VOSviewer*. Recuperado el 20 de octubre de 2018, de <http://www.vosviewer.com>
- Chen, Y., Li, S., Wu, W., Geng, S., y Mao, M. (2022). Distinct mutations and lineages of SARS-CoV-2 virus in the early phase of COVID-19 pandemic and subsequent 1-year global

- expansion. *Journal of Medical Virology*, 94(5), 2035–2049. <https://doi.org/10.1002/jmv.27580>
- Dipta, D., y Naoki, F. (2021). Toward Efficient Analysis of Open Data and Social Phenomenon -a Case Study on Weather Influence on Spreading of SARS-CoV-2. *2021 10th International Congress on Advanced Applied Informatics (IIAI-AAI)*, 903–906. <https://doi.org/10.1109/IIAI-AAI53430.2021.00157>
- Elsevier B.V. (s/f). *Scopus*. Recuperado el 18 de octubre de 2018, de <https://www.scopus.com>
- Eugene, A., Alpert, N., Lieberman-Cribbin, W., y Taioli, E. (2021). Trends in COVID-19 School Related Inquiries Using 311 New York City Open Data. *Journal of Community Health*, 46(6), 1177–1182. <https://doi.org/10.1007/s10900-021-01006-y>
- Francini, S., McRoberts, R. E., D'Amico, G., Coops, N. C., Hermosilla, T., White, J. C., Wulder, M. A., Marchetti, M., Mugnozza, G. S., y Chirici, G. (2022). An open science and open data approach for the statistically robust estimation of forest disturbance areas. *International Journal of Applied Earth Observation and Geoinformation*, 106, 102663. <https://doi.org/10.1016/j.jag.2021.102663>
- Gkiouras, K., Nigdelis, M. P., Grammatikopoulou, M. G., y Goulis, D. G. (2020). Tracing open data in emergencies: The case of the COVID-19 pandemic. *European Journal of Clinical Investigation*, 50(9). <https://doi.org/10.1111/eci.13323>
- Hagen, L., Sandoval-Almazan, R., Okhuijsen, S., Cabaco, S., A. Ruvalcaba-Gomez, E., Villodre, J., Sung, W., y Valle-Cruz, D. (2021). Open Government and Open Data in Times of COVID-19. *DG.O2021: The 22nd Annual International Conference on Digital Government Research*, 598–600. <https://doi.org/10.1145/3463677.3463740>
- Kobayashi, S., Falcón, L., Fraser, H., Braa, J., Amarakoon, P., Marcelo, A., y Paton, C. (2021). Using Open Source, Open Data, and Civic Technology to Address the COVID-19 Pandemic and Infodemic. *Yearbook of Medical Informatics*, 30(01), 038–043. <https://doi.org/10.1055/s-0041-1726488>
- McClary-Gutierrez, J. S., Aanderud, Z. T., Al-faliti, M., Duvallet, C., Gonzalez, R., Guzman, J., Holm, R. H., Jahne, M. A., Kantor, R. S., Katsivelis, P., Kuhn, K. G., Langan, L. M., Mansfeldt, C., McLellan, S. L., Mendoza Grijalva, L. M., Murnane, K. S., Naughton, C. C., Packman, A. I., Paraskevopoulos, S., ... Delgado Vela, J. (2021). Standardizing data reporting in the research community to enhance the utility of open data for SARS-CoV-2 wastewater surveillance. *Environmental Science: Water Research & Technology*, 7(9), 1545–1551. <https://doi.org/10.1039/D1EW00235J>

- Miljenović, D., y Beriša, B. (2022). Pandemics trends in E-commerce: Drop shipping entrepreneurship during COVID-19 pandemic. *Pomorstvo*, 36(1), 31–43. <https://doi.org/10.31217/p.36.1.4>
- Pecoraro, F., y Luzi, D. (2021). Open Data Resources on COVID-19 in Six European Countries: Issues and Opportunities. *International Journal of Environmental Research and Public Health*, 18(19), 10496. <https://doi.org/10.3390/ijerph181910496>
- Petrovich, E. (2022). Bibliometrics in Press. Representations and uses of bibliometric indicators in the Italian daily newspapers. *Scientometrics*, 127(5), 2195–2233. <https://doi.org/10.1007/s11192-022-04341-6>
- Pujo, J. M., Fitriani, D. Y., Adi, N. P., Portecop, P., Resiere, D., Mansyur, M., y Kallel, H. (2022). COVID-19 Outbreak: Burnout and Posttraumatic Stress Disorder, a Harmful Chronology for Health Caregivers in Emergency Departments and Intensive Care Units. *Disaster Medicine and Public Health Preparedness*, 1–2. <https://doi.org/10.1017/dmp.2022.24>
- Simuț, C. C., Petrița, L., Popescu, F.-A., y Oprea, I. M. (2021). Challenges and Opportunities for Telecommuting in the School System: Building a Sustainable Online Education in the Context of the SARS-Cov-2 Pandemic. *Sustainability*, 13(18), 10296. <https://doi.org/10.3390/su131810296>
- Sohel, D., y Md. Rabiul, I. (2022). The SARS-CoV-2 omicron wave is indicating the end of the pandemic phase but the COVID-19 will continue. *Journal of Medical Virology*, 94(6), 2343–2345. <https://doi.org/10.1002/jmv.27635>
- Straub, F., Streppel, S., y Göhlich, D. (2021). Methodology for Estimating the Spatial and Temporal Power Demand of Private Electric Vehicles for an Entire Urban Region Using Open Data. *Energies*, 14(8), 2081. <https://doi.org/10.3390/en14082081>
- Strcic, J., Civljak, A., Glozinic, T., Pacheco, R. L., Brkovic, T., y Puljak, L. (2022). Open data and data sharing in articles about COVID-19 published in preprint servers medRxiv and bioRxiv. *Scientometrics*, 127(5), 2791–2802. <https://doi.org/10.1007/s11192-022-04346-1>
- van Eck, N. J., y Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>
- van Eck, N. J., y Waltman, L. (2014). Visualizing Bibliometric Networks. En Y. Ding, R. Rousseau, & D. Wolfram (Eds.), *Measuring Scholarly Impact* (pp. 285–320). Springer International Publishing. https://doi.org/10.1007/978-3-319-10377-8_13
- Wahlteinez, O., Cheung, A., Alcantara, R., Cheung, D., Daswani, M., Erlinger, A., Lee, M., Yawalkar, P., Lê, P., Navarro, O. P., Brenner, M. P., y Murphy, K. (2022). COVID-19 Open-Data a

global-scale spatially granular meta-dataset for coronavirus disease. *Scientific Data*, 9(1), 162. <https://doi.org/10.1038/s41597-022-01263-z>

Wardle, J. A., Sagan, V., y Mohammed, F. (2022). Using open data cube on the cloud to investigate food security by means of cropland changes in djibouti. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B3-2022, 1039–1044. <https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-1039-2022>