

Análisis bibliométrico con minería de texto sobre *helicobacter pylori* y plasma rico en plaquetas

Bibliometric Analysis with Text Mining on Helicobacter Pylori and Platelet-Rich Plasma

José Aureliano Betancourt Bethencourt^{1*} <https://orcid.org/0000-0003-0043-9526>

Elizabeth Nicolau Pestana¹ <https://orcid.org/0000-0003-1975-3419>

Lidyce Quesada Leyva¹ <https://orcid.org/0000-0002-4250-0164>

¹Universidad de Ciencias Médicas de Camagüey, Centro de Inmunología y Productos Biológicos (CENIPBI). Camagüey, Cuba.

*Autor para la correspondencia: betanster@gmail.com

RESUMEN

Introducción: La investigación en salud ha experimentado un crecimiento exponencial en los últimos años, impulsada por avances tecnológicos y el acceso a una gran cantidad de información científica; sin embargo, la minería de texto, una disciplina que permite extraer información valiosa a partir de grandes volúmenes de texto, no se emplea sistemáticamente.

Objetivo: Analizar las publicaciones y los temas investigados sobre *Helicobacter pylori* y plasma rico en plaquetas con la técnica de minería de texto.

Métodos: Se accedió a las bases de datos de Europa PMC mediante el programa R en el que se solicitó la cantidad de las publicaciones y los temas investigados sobre *helicobacter pylori* y plasma rico en plaquetas, durante el período 2010-2022. Se trabajó con 10 000 artículos por cada tema. Con el paquete tm se logró construir una nube de palabras que fue

presentada a los profesores e investigadores. Los textos con las informaciones aportadas por los investigadores se procesaron y analizaron.

Resultados: Se evaluaron 20 000 publicaciones y los temas investigados sobre *helicobacter pylori* y plasma rico en plaquetas. Se generaron dos nubes de palabras; con ellas se estableció un debate y una tormenta de ideas.

Conclusiones: La minería de texto fue de gran provecho para el análisis de los temas de *helicobacter pylori* y el plasma rico en plaquetas. Se realizó un análisis con sus beneficios para las actividades cognitivas y creativas de los investigadores. Se identificaron nuevos temas y las relaciones entre ellos.

Palabras clave: minería de textos; *helicobacter pylori*; investigadores; plasma rico en plaquetas; ciencia de datos.

ABSTRACT

Introduction: Health research has experienced exponential growth in recent years, driven by technological advances and access to large amounts of scientific information; however, text mining, a discipline that allows extracting valuable information from large volumes of text, is not used systematically.

Objective: To analyze the publications and research topics on *Helicobacter pylori* and platelet-rich plasma using the text mining technique.

Methods: The databases of Europe PMC were accessed through the R program, in which the number of publications and the topics investigated on *helicobacter pylori* and platelet-rich plasma were requested, from 2010 to 2022. We worked with 10,000 articles for each topic. With the tm package it was possible to build a word cloud that was presented to professors and researchers. The texts with the information provided by the researchers were processed and analyzed.

Results: 20,000 publications and research topics on *helicobacter pylori* and platelet-rich plasma were evaluated. Two-word clouds were generated, a debate and a brainstorming session were established with them.

Conclusions: Text mining was of great benefit for the analysis of *helicobacter pylori* and platelet-rich plasma issues. An analysis was carried out with its benefits for the cognitive

and creative activities of the researchers. New themes and relationships between them were identified.

Keywords: text mining; *helicobacter pylori*; researchers; platelet-rich plasma; data science.

Recibido: 09/11/2022

Aceptado: 08/05/2023

Introducción

La investigación en salud ha experimentado un crecimiento exponencial en los últimos años, impulsada por avances tecnológicos y el acceso a una gran cantidad de información científica. Sin embargo, la explotación sistemática de los datos textuales en esta área aún presenta desafíos. En particular, la minería de texto, una disciplina que permite extraer información valiosa a partir de grandes volúmenes de texto, no se utiliza de manera generalizada en la investigación en salud.⁽¹⁾

Algunas de las técnicas utilizadas en la minería de texto en el ámbito de la salud pueden ayudar a los investigadores a obtener información valiosa, identificar patrones y tendencias y descubrir temas emergentes que pueden guiar futuras investigaciones y mejorar la toma de decisiones en salud. Algunas de las más utilizadas incluyen:⁽²⁾

1. Análisis de frecuencia de palabras: Consiste en identificar las palabras más frecuentes en un conjunto de textos relacionados con la salud. Esto puede ayudar a destacar los temas principales que se abordan en la investigación y detectar las palabras clave más relevantes.
2. Análisis de co-ocurrencia: Esta técnica busca identificar las palabras que tienden a aparecer juntas con mayor frecuencia. Por ejemplo, se puede analizar qué términos

- suelen acompañar a una enfermedad específica, lo que revela patrones de síntomas, tratamientos o factores de riesgo asociados.
3. **Análisis de sentimiento:** En el ámbito de la salud, el análisis de sentimiento se utiliza para evaluar las opiniones, actitudes y emociones expresadas en textos, como reseñas de pacientes, comentarios en redes sociales o informes de estudios clínicos. Esto puede ayudar a identificar patrones de satisfacción, preocupación o efectividad en relación a tratamientos o intervenciones específicas.
 4. **Modelos de temas:** Los modelos de temas son algoritmos que agrupan automáticamente los documentos en categorías o temas basados en el contenido textual. Estos modelos pueden identificar temas emergentes en la investigación en salud, lo que permite a los investigadores estar al tanto de las últimas tendencias y áreas de interés.
 5. **Análisis de redes:** El análisis de redes se utiliza para examinar las relaciones entre conceptos o entidades en un conjunto de textos. En el contexto de la salud esto puede contribuir a identificar las conexiones entre enfermedades, tratamientos, genes, proteínas u otros elementos relevantes. También se puede utilizar para analizar redes de colaboración entre investigadores y cómo se relacionan entre sí.

El análisis de grandes cantidades de datos textuales requiere de un trabajo arduo para leer todo el documento y organizar el contenido, por ende, su procesamiento es cuestionable. La minería de textos permite extraer información y conocimiento interesantes y no triviales de un texto no estructurado, así como detectar patrones significativos implícitamente presentes en los datos.⁽³⁾

La nube de palabras derivadas de la minería de texto es una herramienta de visualización de contenidos que facilita al investigador realizar análisis descriptivos de un documento. Este tipo de visualización resalta la frecuencia con que se utilizan los términos u otro análisis asociado a las palabras. La frecuencia de las palabras en los textos denota la importancia de ellas en cada contexto, por lo que cada palabra puede llegar a capturar el propósito e ideas de los autores.⁽⁴⁾

La nube de palabras o nube de etiquetas o de texto es una representación visual de la frecuencia y el valor de las palabras. Con ella se resalta la frecuencia con la que un término específico o categoría aparece en una fuente de datos. Las palabras que aparecen con más frecuencia en un texto se muestran en un tamaño mayor. Representan jerárquicamente los

términos presentados con mayor frecuencia, lo que permite agilizar el análisis de datos tipo texto, puesto que, a través de ellos, se puede identificar e interpretar de manera rápida y sencilla las palabras con mayor relevancia en el texto analizado, ayudan a trabajar la capacidad de síntesis, fomentan la creatividad y contribuyen a establecer el índice de los temas que se van a tratar o a modo de resumen de cada apartado que se aborde.^(5,6)

En el Centro de Inmunología y Productos Biológicos de Camagüey se trabajan dos temas: uno es sobre el *Helicobacter pylori*,⁽⁷⁾ en particular los medios diagnósticos como el test de ureasa,⁽⁸⁾ y el otro el plasma rico en plaquetas sobre todo su preparación.⁽⁹⁾ Por ello es que se exploran las publicaciones sobre estos dos temas.

Helicobacter pylori es un patógeno emergente en el ser humano que provoca dispepsias, úlceras y otros trastornos.⁽¹⁰⁾ El plasma rico en plaquetas es un producto derivado de la sangre, rico en péptidos y proteínas de señalización intercelular, así como citoquinas, capaces de intervenir en cada una de las etapas de la regeneración de varios tejidos. Principalmente, se le han atribuido efectos antiinflamatorios en diferentes lesiones, así como otros biológicos sobre las células y tejidos.⁽¹¹⁾

En este contexto, la investigación tuvo como objetivo analizar las publicaciones y temas investigados acerca del *Helicobacter pylori* y el plasma rico en plaquetas con la técnica de minería de texto.

Métodos

Se accedió a las bases de datos de Europa PMC,^(12,13) mediante el programa R⁽¹⁴⁾ en el que se solicitaron 20 000 publicaciones sobre a) *helicobacter pylori* b) plasma rico en plaquetas (PRP), durante el período del 2010 al 2022. Se limitaron los artículos a trabajar a 10 000 por cada tema, como muestra intencional.

Con el paquete tm,⁽¹⁵⁾ que tiene métodos para la importación de datos, manejo de corpus, preprocesamiento, gestión de metadatos y creación de matrices término-documento, se preparó el documento para su análisis, se convirtió el formato de texto a documento, el cual se inspeccionó. Se eliminaron palabras y puntuaciones no necesarias, se preparó la matriz

del texto, se buscaron los elementos más frecuentes y se trazó un gráfico de nube de palabras con el paquete worcloud2.⁽⁶⁾

Estas dos salidas del procesamiento anterior sobre *helicobacter pylori* y plasma rico en plaquetas (PRP) se expusieron para debates, mediante una tormenta de ideas. A los profesores e investigadores del Centro de Inmunología y Productos Biológicos (CENIPBI) de la Universidad de Ciencias Médicas de Camagüey se les pidió que refirieran, a su juicio, en un documento de texto si les parecía útil y por qué; qué tema se destaca y cómo lo investigaría, la relación con el centro que observa en cada tema y algún aspecto que les haya llamado la atención. Los textos con estas informaciones aportadas por el conjunto de investigadores se procesaron con el paquete tm, mediante el procedimiento descrito; se preparó una nube de palabras con esta información con la que se llevó a efecto un debate y una tormenta de ideas.

Resultados

Se lograron evaluar 20 000 publicaciones y los temas investigados sobre *helicobacter pylori* y plasma rico en plaquetas, durante el período del 2010 al 2022.

Las frecuencias de publicaciones para ambos aspectos tienen un crecimiento exponencial, lo que indica el gran interés en estos temas (fig. 1).

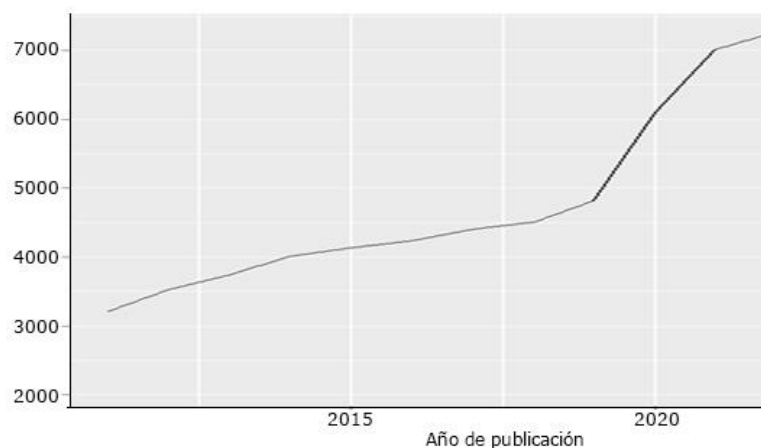


Fig. 1 – Publicaciones por año.

La salida del programa R tiene concebido el conteo de palabras de los aspectos publicados. En este caso se extrajo la de *Helicobacter pylori*. Esta es una manera de valorar en un histograma (fig. 2) las frecuencias de determinada cantidad de palabras lo que permite valorar la importancia de las palabras utilizadas. Los histogramas son una herramienta gráfica utilizada para representar la distribución de una variable. En general, la forma del histograma puede indicar el tipo de distribución de la variable. Por ejemplo, si la distribución es simétrica y en forma de campana, se dice que sigue una distribución normal, lo cual no es así en este caso donde la distribución es asimétrica.⁽¹⁶⁾

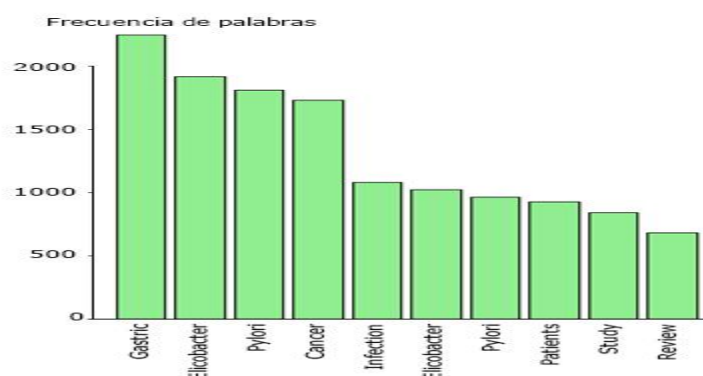


Fig. 2 – Palabras más frecuentes para *helicobacter pylori*.

Una nube de palabras para evaluar las publicaciones científicas en el campo de la salud podría contener términos como calidad, evaluación, metodología, control, criterios, resumen, palabras clave, estructura, claridad, precisión, indicadores, proceso editorial, calidad científica, calidad informativa, entre otros. Además, la evaluación de la calidad de las publicaciones científicas en el campo de la salud puede incluir aspectos como la coincidencia entre palabras clave y descriptores, el uso de aplicaciones móviles con respaldo en las publicaciones científicas y la propuesta de indicadores de evaluación del funcionamiento y la calidad de las revistas científicas en Ciencias de la Salud.⁽¹⁷⁾

En este trabajo se generaron dos nubes de palabras de las bibliografías (fig. 3 y 4) y una del debate realizado con los investigadores (fig. 5).

En la observación de las nubes de palabras se destaca lo siguiente:

- a) Estas salidas pueden propiciar debates y llevar a efecto la tormenta de ideas.
- b) Sería interesante estudiar qué relación tienen los diabéticos con el PRP, a partir de la información gráfica.
- c) La Alta frecuencia con la que se cita el término *prion* que condujo a indagar qué relación tiene con el PRP.
- d) Identificar posibles relaciones entre variables pobremente estudiadas.
- e) Se realizan revisiones sistémicas y metaanálisis de estos temas que pueden consultarse.
- f) La posibilidad de encaminar tratamientos alternativos.
- g) El uso de experimentos con animales de laboratorio y otros modelos.
- h) La nube de palabras permite reducir el tiempo que se le dedica a la lectura de textos extensos en función de las búsquedas; algunos investigadores recomiendan analizar aun a las palabras pequeñas.
- i) La población local es envejecida y tiene ambas afecciones.

Discusión

A través de la generación de nubes de palabras y la participación activa de expertos, se logró identificar nuevas perspectivas y relaciones entre los temas de interés. Estos hallazgos representan un avance significativo en el análisis de datos textuales en la investigación en salud y brindan beneficios, tanto en términos de actividades cognitivas como creativas para los investigadores.

Este enfoque ha permitido el análisis exhaustivo de temas relacionados con *Helicobacter pylori* y el plasma rico en plaquetas; identificar nuevas áreas de investigación y fomentar la generación de ideas innovadoras. A través de este estudio, se espera impulsar el uso sistemático de la minería de texto en la investigación en salud y promover la colaboración interdisciplinaria para abordar los desafíos científicos en esta área.

La minería de texto y el análisis bibliométrico realizado en este estudio han permitido obtener información valiosa sobre los temas de *Helicobacter pylori* y el plasma rico en

plaquetas (PRP) en la investigación en salud. Los resultados revelaron un crecimiento exponencial en la cantidad de publicaciones sobre ambos temas durante el período 2010 al 2022, lo que indica un gran interés en estas áreas de estudio.

El análisis de frecuencia de palabras y la generación de nubes de palabras han proporcionado una visión general de los términos más relevantes y frecuentes asociados con *Helicobacter pylori* y PRP. Estas visualizaciones han facilitado la identificación de temas principales y han abierto la puerta a debates y tormentas de ideas entre investigadores y expertos en el Centro de Inmunología y Productos Biológicos de la Universidad de Ciencias Médicas de Camagüey.

Dentro de las observaciones destacadas se encuentra la alta frecuencia de términos como “diabéticos” y “prion” en relación con el PRP, lo que sugiere posibles relaciones y áreas de investigación interesantes. Además, se ha identificado la necesidad de investigar variables pobremente estudiadas y explorar tratamientos alternativos. También se señala la presencia de revisiones sistémicas y metaanálisis sobre estos temas, lo que proporciona una base sólida para consultas posteriores.

El uso de modelos de temas y análisis de redes ha permitido identificar nuevas perspectivas y relaciones entre los temas de interés. Estos hallazgos abren oportunidades para la generación de ideas innovadoras y fomentan la colaboración interdisciplinaria en la investigación en salud.

Es importante destacar que el enfoque de minería de texto y análisis bibliométrico utilizado en este estudio tiene implicaciones prácticas y científicas significativas. En primer lugar, permite una exploración exhaustiva de los temas de investigación en salud, lo que puede llevar a una mejor comprensión de los avances y desafíos en el campo. Además, fomenta la colaboración entre investigadores, al facilitar la identificación de áreas de interés común y promover la generación de ideas.

Sin embargo, es necesario mencionar algunas limitaciones. El análisis bibliométrico se basa en los datos disponibles en las bases de datos seleccionadas, lo que puede generar sesgos y no reflejar completamente todas las investigaciones realizadas en el campo. Además, el procesamiento del texto y la generación de nubes de palabras dependen de las técnicas y algoritmos utilizados, lo que puede influir en los resultados obtenidos.

En conclusión, la minería de texto y el análisis bibliométrico son herramientas poderosas en la investigación en salud que permiten extraer información valiosa y descubrir patrones en grandes volúmenes de texto. En este estudio, se ha demostrado su utilidad en el análisis de los temas de *Helicobacter pylori* y plasma rico en plaquetas.

Zhang y otros en su estudio emplearon técnicas de minería de texto y modelación para extraer los temas de investigación de pesquisas públicas; se demuestra la utilidad de estas técnicas asociadas a las redes en la salud pública y recomiendan estudios posteriores.⁽¹⁸⁾ Ford y otros⁽¹⁹⁾ encontraron que las técnicas de minería de texto están avanzando cada vez más y recomiendan incrementar los estándares éticos. Chandril y otros argumentan que el empleo de narrativas son una fuente de datos estadísticamente viables que permite distinguir entre pacientes con diferentes diagnósticos. Rizk y otros⁽²⁰⁾ muestran un ejemplo para estudiar la innovación digital, como guía a los investigadores en el uso de análisis de textos. Se realizan esfuerzos para fomentar la minería de textos en biomedicina.⁽²¹⁾

En general, este estudio demuestra que la minería de texto con el programa R fue de gran utilidad para los investigadores del centro estudiado, especialmente en el análisis de los temas de *Helicobacter pylori* y el plasma rico en plaquetas. Se lograron identificar nuevas perspectivas y relaciones entre los temas de interés, lo cual puede impulsar el uso sistemático de la minería de texto en la investigación en salud y promover la colaboración interdisciplinaria

Conclusiones

En este estudio se logró realizar una evaluación exhaustiva de 20 000 publicaciones relacionadas con los temas de *Helicobacter pylori* y plasma rico en plaquetas durante el período comprendido entre 2010 y 2022. A través del análisis de los datos, se generaron dos nubes de palabras que destacan las palabras más frecuentes asociadas con cada tema.

La discusión general y los debates generados, a partir de las nubes de palabras, revelaron varias conclusiones importantes. En primer lugar, se reconoció que la metodología de la minería de texto puede facilitar debates y tormentas de ideas, lo que fomenta la generación de nuevas perspectivas e ideas innovadoras. Además, surgió el interés de investigar la relación entre PRP y pacientes diabéticos, así como la frecuente aparición del término “prion” en relación con PRP, lo que motivó una indagación adicional sobre su posible relación.

También se identificó la necesidad de explorar las relaciones entre variables poco estudiadas y la realización de revisiones sistemáticas y metaanálisis en relación con estos temas. Se destacó la posibilidad de dirigir tratamientos alternativos basados en los hallazgos obtenidos, así como la importancia de utilizar modelos en animales de laboratorio y otros enfoques experimentales.

Se resaltó el valor de las nubes de palabras como herramienta para reducir el tiempo dedicado a la lectura de textos extensos y facilitar las búsquedas, incluso, recomendando el análisis de palabras más pequeñas.

Además, se observó que la población local estudiada en las publicaciones era envejecida y presentaba ambas afecciones, lo que proporciona información importante sobre el contexto demográfico y clínico de los estudios.

En conclusión, este estudio demuestra la utilidad de la minería de texto en el análisis de investigaciones en salud, específicamente en los temas de *Helicobacter pylori* y plasma rico en plaquetas. Los resultados obtenidos permitieron identificar nuevos temas de investigación, establecer relaciones entre variables poco exploradas y generar debates y tormentas de ideas en el ámbito científico. Estas conclusiones respaldan estudios previos que resaltan el potencial de la minería de texto en el campo de la salud pública y enfatizan la necesidad de incrementar los estándares éticos en su aplicación. Este enfoque de análisis de textos ofrece una guía útil para los investigadores interesados en explorar la innovación digital y el uso de técnicas de minería de texto en sus estudios.

Referencias bibliográficas

1. Lemus-Delgado D, Pérez Navarro R. Ciencia de datos y estudios globales: aportaciones y desafíos metodológicos. Colombia Internacional. 2020 [12/10/2022];(102):41-62. Disponible en: <https://revistas.uniandes.edu.co/index.php/colombia-int/article/view/2923>
2. Mariñelarena-Dondena L, Errecalde ML, Castro Solano A. Extracción de conocimiento con técnicas de minería de textos aplicadas a la psicología. Revista Argentina de Ciencias del Comportamiento. 2017 [05/10/2022];9(2):0. Disponible en: http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1852-42062017000200006

3. Hassani H, Beneki C, Unger S, Mazinani MT, Yeganegi MR. Text mining in big data analytics. *Big Data and Cognitive Computing*. 2020 [12/10/2022];4(1):1p. Disponible en: <https://www.mdpi.com/2504-2289/4/1/1>
4. Lee WJ. A study on word cloud techniques for analysis of unstructured text data. *The Journal of the Convergence on Culture Technology*. 2020 [19/10/2022];6(4):715-20. Disponible en: <https://www.koreascience.or.kr/article/JAKO202034965718309.page>
5. Holtz Y. The Wordcloud2 Library. Retrieved from the R Graph Gallery. 2018. Disponible en: <https://r-graph-gallery.com/196-the-wordcloud2-library.html>
6. Lang DCG. wordcloud2: Create Word Cloud by 'htmlwidget'. R package version 0.2.12018. Disponible en: <https://CRAN.R-project.org/package=wordcloud2> .
7. Morales Díaz M, Corrales Alonso S, Vanterpoll Héctor M, Avalos Rodríguez R, Salabert Tortolo I, Hernández Diaz O. Cáncer gástrico: algunas consideraciones sobre factores de riesgo y Helicobacter pylori. *Revista médica electrónica*. 2018 [19/10/2022];40(2):433-44. Disponible en: www.scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1684-18242018000200018
8. Graham DY, Miftahussurur M. Helicobacter pylori urease for diagnosis of Helicobacter pylori infection: A mini review. *Journal of advanced research*. 2018 [19/10/2022];13:51-7. Disponible en: <https://www.sciencedirect.com/science/article/pii/S209012321830016X> .
9. Aguilar R, Cáceres A. Plasma rico en plaquetas como terapia autóloga en la medicina regenerativa: Revisión narrativa. *Ciencia, Tecnología y Salud*. 2020 [19/10/2022];7(3):442-60. Disponible en: <https://revistas.usac.edu.gt/index.php/cytes/article/view/946>
10. Pérez C, Escobar G, Soto P, Ibarra C, Larrondo M, Alfaro J. Obtención de plasma rico en plaquetas (PRP) en el Laboratorio de Terapia Celular para uso como herramienta terapéutica en medicina regenerativa. *Rev Hosp Clin Univ Chile*. 2021 [19/10/2022]:4-16. Disponible en: <https://www.redclinica.cl/Portals/0/Users/014/14/14/1950.pdf>
11. Jiménez Jiménez G. Helicobacter pylori como patógeno emergente en el ser humano. *Revista Costarricense de Salud Pública*. 2018 [19/10/2022];27(1):65-78. Disponible en: https://www.scielo.sa.cr/scielo.php?pid=S1409-14292018000100065&script=sci_arttext

12. Europe PMC Consortium. Europe PMC: a full-text literature database for the life sciences and platform for innovation. *Nucleic acids research*. 2015;43(D1):D1042-8. DOI: <https://doi.org/10.1093/nar/gku1061>
13. PMC. Europe PMC 2022. Disponible en: <https://europepmc.org/>
14. Valero Moreno AI. Técnicas estadísticas en Minería de Textos. (Tesis de fin de grado). Universidad de Sevilla, Sevilla. 2017 [19/10/2022]. Disponible en: <https://idus.us.es/handle/11441/63197>
15. Feinerer I. Introduction to the tm Package Text Mining in R. 2022 [19/10/2022]. Disponible en: <https://cran.r-project.org/web/packages/tm/vignettes/tm.pdf> .
16. Gutiérrez RB, i Cintas PG. El histograma como un instrumento para la comprensión de las funciones de densidad de probabilidad. *Probabilidad Condicionada: Revista de didáctica de la Estadística*. 2013 [19/10/2022];(2):229-35. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=4770286>
17. De Lucia Castillo F, Saibel Santos CA. Nubes de palabras animadas para la visualización de información textual de Publicaciones Académicas. Universidad Garcilaso de la Vega. Repositorio Institucional. 2016. Disponible en: <http://repositorio.uigv.edu.pe/handle/20.500.11818/615> .
18. Zhang Y, Cao B, Wang Y, Peng TQ, Wang X. When public Health Research meets social media: knowledge mapping from 2000 to 2018. *Journal of medical Internet research*. 2020 [19/10/2022];22(8):e17582. Disponible en: <https://www.jmir.org/2020/8/e17582>
19. Ford E, Shepherd S, Jones K, Hassan L. Toward an ethical framework for the text mining of social media for health research: a systematic review. *Frontiers in Digital Health*. 2021 [19/10/2022];2:592237. Disponible en: <https://www.frontiersin.org/articles/10.3389/fdgth.2020.592237/full> .
20. Rizk A, Elragal A. Data science: developing theoretical contributions in information systems via text analytics. *Journal of Big Data*. 2020 [19/10/2022];7:1-26. Disponible en: <https://link.springer.com/article/10.1186/s40537-019-0280-6> .
21. Villegas M, de la Peña S, Intxaurreondo A, Santamaria J, Krallinger M. Esfuerzos para fomentar la minería de textos en biomedicina más allá del inglés: el plan estratégico nacional español para las tecnologías del lenguaje. *Procesamiento del Lenguaje Natural*. 2017

[19/10/2022];(59):141-4.

Disponible

en:

<https://www.redalyc.org/pdf/5157/515754427019.pdf>

Conflicto de intereses

Los autores declaran que no tienen conflicto de intereses.

Contribuciones de los autores

Conceptualización: José Aureliano Betancourt Bethencourt, Elizabeth Nicolau Pestana, Lidyce Quesada Leyva.

Curación de datos: José Aureliano Betancourt Bethencourt.

Análisis formal: José Aureliano Betancourt Bethencourt.

Software: José Aureliano Betancourt Bethencourt.

Investigación: Elizabeth Nicolau Pestana, Lidyce Quesada Leyva.

Metodología: José Aureliano Betancourt Bethencourt.

Redacción – borrador original: José Aureliano Betancourt Bethencourt.

Redacción – revisión y edición: Elizabeth Nicolau Pestana, Lidyce Quesada Leyva.