

Análisis Híbrido de la producción académica: Bibliometría y Revisión Sistemática asistida por Machine Learning

Carlos Schenone

2023-06-23

Contents

Resumen de Títulos	2
Título “Análisis Híbrido: Análisis bibliométrico y Revisión Sistemática asistida por Machine Learning”	3
Resumen	3
Palabras clave	3
Abstract	3
KeyWords	3
Introducción	4
Comentar los puntos de investigación (ver el artículo de bariviera and merediz sobre sobre análisis híbrido en investigaciones sobre blockchain)	4
Justificación de la investigación	4
Objetivos	5
Planteamiento del problema	5
Descripción de resto del documento	6
Fundamentos teóricos de la investigación	6
Marco conceptual	6
Marco Teórico	6
Hipótesis y Variables	9
Formulación de hipótesis	9
Identificación de Variables	9
Operacionalización de variables	11

Metodología	11
Diseño y Tipo del Estudio	11
Población	11
Muestra	12
Flujograma de la Metodología	12
Técnicas y procedimiento	13
Análisis de resultados	14
Resultados	18
Síntesis de los resultados	18
Plan de análisis	18
Análisis de resultados	21
Consideraciones éticas	25
Potenciales líneas de investigación	25
Conclusiones	25
Conclusiones	25
Expresiones de gratitud	26
Referencias o Referencias bibliográficas	26
Bibliografía	26
Apéndices	26
–Retomar desde la INTRODUCCIÓN	

Resumen de Títulos

Título Resumen Palabras clave Introducción - Justificación - Objetivos - Planteamiento del problema - Descripción del resto del documento Fundamentos teóricos de la investigación (marco teórico) - Marco conceptual - Marco teórico Hipótesis y variables (va?) Metodología - Diseño y tipo de estudio - Población - Muestra – Criterios de inclusión y exclusión - Técnicas y procedimiento – Descripción del laboratorio – Selección de la BD a consultar – Definición y análisis de los metadatos de interés – Descripción de los criterios de extracción, transformación y carga – Validación de performance – Flujograma de tamizaje de artículos Análisis de resultados - Indicadores bibliométricos - Análisis de las visualizaciones Resultados Consideraciones éticas Potenciales líneas de investigación Conclusiones Expresiones de gratitud Referencias o Referencias bibliográficas Apéndices

Nota: Según PRISMA 2020.

Título “Análisis Híbrido: Análisis bibliométrico y Revisión Sistemática asistida por Machine Learning”

Identifique el informe o publicación como una revisión sistemática. (en mi caso es un análisis bibliométrico)

Resumen

Insertar acá un breve resumen (menos de 2.000 palabras) del tema del TP final.

Análisis bibliométrico o bibliometría Definición (BUSCAR DEFINICIÓN EN UN LIBRO O EN EL ARTICULO DEL PROFESOR DEL CURSO SOBRE BITCOIN. ESTE LO SAQUE DE OTRO ARTICULO PERO ES UN COMENTARIO: La bibliometría evalúa la calidad y cantidad de la literatura científica publicada, colaborando en la medición de la productividad de la investigación)

Revisión sistemática Definición: ...

Análisis híbrido Definición: ...

Machine learningng aplicado al análisis de texto (1. BUSCAR DEFINICION DE ML aplicado a Text Minning EN UN ARTICULO O LIBRO) Definición: ...

LUEGO SIGUE ESTO:

La producción científica de artículos sobre inteligencia artificial aplicada al área de análisis de videos fue descrita por (# de artículos) de 2006 a 2022, por lo que se desconoce la realidad de la producción científica en esta área específica hasta el día de hoy. En este sentido, el artículo se propone analizar y describir la evidencia científica (calidad y cantidad) sobre inteligencia artificial aplicada al área de análisis de videos. Para abordar el problema, se diseñó un estudio siguiendo las recomendaciones del análisis bibliométrico (buscar un artículo que recomienda los pasos a seguir) tomando como insumo los artículos científicos sobre el tema de interés entre hasta diciembre 2022, publicados en revistas indexadas sistematizadas en la base de datos Dimensions. Como herramienta se utiliza el software R, R Studio y la biblioteca bibliometrix (buscar artículo donde se proponga el uso de R y Bibliomwtrix). Como resultado de aplicar el filtro xxxxx se obtuvieron xxxxxx citas en la base de datos. Se examinaron para evaluar si cumplían con los criterios de selección, excluyendo xxxx artículos, obteniendo finalmente xxxx artículos incluidos en el presente análisis. Como conclusión podemos mencionar (destacar los principales puntos devueltos por bibliometrix, mencionar el crecimiento exponencial de las publicaciones, y en detalle cuáles países lideran el ranking)

Palabras clave

Análisis bibliométrico, análisis de la producción científica, revisión sistemática, inteligencia artificial, machine learning.

Abstract

Insertar acá la traducción al inglés del Resumen (menos de 2000 palabras).

KeyWords

Insertar acá la traducción al inglés de las palabras clave

Introducción

La introducción consiste en una descripción clara y precisa del problema que se investigó. En ella se especifican el tema de estudio, los antecedentes, la justificación, el problema, la hipótesis y los objetivos. Los antecedentes y la justificación se refieren a la parte donde se mencionan tanto la necesidad como la utilidad de llevar a cabo la investigación propuesta. Se debe incluir una descripción muy breve de los hallazgos más sobresalientes en relación con el tema de interés, especificando la población en la cual se efectuaron dichas investigaciones. El problema planteado debe ser original, relevante y verificable (Ibáñez Brambilla, 151). La hipótesis sirve como una guía que orienta la ruta que debe seguirse para la obtención de datos, y es la respuesta tentativa que se propone al problema de investigación. Finalmente, los objetivos se refieren a los logros que se buscan obtener al final del trabajo y son justamente los que determinan si el proyecto resultó como se esperaba o no.

En la introducción también se integra una descripción de cada uno de los capítulos que forman parte del trabajo. Es decir, se hace una breve sinopsis del contenido de cada uno de los apartados que conforman la tesis.

Comentar los puntos de investigación (ver el artículo de bariviera and merediz sobre sobre análisis híbrido en investigaciones sobre blockchain)

La tecnología blockchain, y más concretamente bitcoin (una de sus principales aplicaciones), ha estado recibiendo una atención creciente en la comunidad científica. Las primeras publicaciones con bitcoin como tema, se remonta a 2012. A pesar de este corto período de tiempo, la producción magnitud (1162 artículos) obliga a realizar un estudio bibliométrico para observar grupos de investigación, temas emergentes y académicos destacados. Nuestro trabajo tiene como objetivo estudiar la producción científica solo en torno a bitcoin, excluyendo otras aplicaciones de blockchain. Por lo tanto, nosotros restringimos nuestra búsqueda a artículos indexados en Web of Science Core Collection, cuyo tema es “bitcoins”. Esta base de datos es adecuada para disciplinas tan diversas como economía, ingeniería, matemáticas y ciencias de la computación. Este estudio bibliométrico dibuja el panorama de la actualidad estado y tendencias de la investigación relacionada con bitcoin en diferentes disciplinas científicas.

Justificación de la investigación

Sobre el objeto de estudio

La finalidad e importancia de un proyecto responde a la pregunta ¿para qué se hace o, qué se quiere hacer con los resultados que se obtengan? y está muy relacionado con el impacto que se espera de los resultados.

Sin embargo, la importancia de los estudios bibliométricos va más allá del ámbito institucional. Podrían ayudar a los nuevos investigadores de una disciplina para comprender el alcance de un tema, las tendencias emergentes y su evolución a lo largo del tiempo. En este sentido es diferente de un estudio de la literatura tradicional. Este tipo de análisis es posible gracias a la disponibilidad de grandes bases de datos como Web of Science.

En consecuencia, nuestro artículo podría considerarse una contribución ampliada a la literatura, proporcionando una visión general completa de la actual tendencias en la investigación de bitcoin e identificación de los mejores investigadores, instituciones y revistas en este campo.

Sobre la originalidad del estudio

Contraste con paper similares La investigación publicada sobre bitcoin comenzó en 2012, y hasta enero de 2019 hay 1162 artículos sobre este tema. Este papel en algunos aspectos, amplía el de Holub y Johnson (2018), añadiendo datos de casi dos años. A diferencia de Holub y Johnson (2018), que utiliza múltiples fuentes de datos, restringimos nuestra búsqueda a Web of Science. El motivo de tal restricción es que Web of Science es el más importante base de datos de artículos científicos, cuyas métricas son ampliamente utilizadas en la evaluación académica en varios países. Además, mejora el análisis agregando representación gráfica de citas, revistas y redes de autores, utilizando el software VOSviewer.

Este trabajo realiza un estudio analítico de los artículos incluidos en la muestra utilizando el paquete R “bibliometrix” (Cuccurullo et al., 2016; Aria y Cuccurullo, 2017). Además, añadimos un análisis gráfico del material bibliográfico utilizando el software VOSviewer (van Eck y Waltman, 2010). Este software recopila los datos y genera mapas basados en acoplamiento bibliográfico, coautoría, citación, coautoría, citación y co-ocurrencia de palabras clave (Merigó et al., 2016).

Objetivos

Objetivo General

- Describir la evidencia científica existente sobre aplicativos móviles en salud en Latinoamérica.

El objetivo del documento, es llevar adelante un análisis bibliométrico de la producción científica de inteligencia artificial aplicada al análisis de videos”. Se plantea el uso de las técnicas de bibliometría para abordar la revisión sistemática debido a la gran cantidad de artículos.

Objetivos del paper (Objetivos: Según PRISMA 2020 “Proporcione una declaración explícita de los principales objetivos o preguntas que aborda la revisión”).

El objetivo de este artículo es analizar los metadatos de todos los artículos indexados en Web of Science Core Collection, cuyo tema es “bitcoin”, excluyendo otras aplicaciones de blockchain. Este estudio está dirigido a un público amplio y diverso. Las partes interesadas de bitcoin vienen de varios campos: programadores informáticos, inversores, pensadores libertarios y economistas financieros. Este documento proporciona útiles información para ellos sobre las principales revistas interesadas en publicar artículos sobre este tema, así como sobre la evolución de los temas abordados en esos papeles. Además, discutimos otros aspectos, como artículos muy citados, principales fuentes de publicación y análisis de palabras clave.

Objetivos específicos

- Estimar la distribución por año de artículos científicos sobre aplicativos móviles en salud en Latinoamérica.
- Categorizar la distribución por autor y país de artículos científicos sobre aplicativos móviles en salud en Latinoamérica.
- Identificar las redes de colaboración entre países que publican artículos científicos sobre aplicativos móviles en salud en Latinoamérica.

Planteamiento del problema

fuente: Crowther M, Lim W, Crowther MA. Systematic Review and meta-analysys methodology. Blood, The Journal of the American society of Hematology. 2010. Oct.

¿Cuál es la producción científica acerca de los aplicativos móviles en salud en Latinoamérica?

Descripción de la fuente de datos

Búsqueda en una base de datos - EB (¿Dónde buscar? Depende del área temática, en general Scopus o WoS) Descripción de las BD a analizar. Descripción de las bases de datos disponibles. Tomar de la página de las BDs. Plantear la descripción de los siguientes atributos de las fuentes de datos: Data Sources / Data Frequency / Main Research Topics

Descripción de los Filtros en la búsqueda

Se pueden definir varios criterios y luego se los suman, restan o se pide que aplique a ambos criterios

Descripción del proceso de extracción de datos

Descripción del proceso de Análisis a aplicar a los datos

Descripción de resto del documento

El resto del documento está estructurado de la siguiente manera. - La sección 2 describe la literatura sobre estudios bibliométricos en revistas económicas. - La sección 3 detalla los datos bajo análisis y comenta los principales hallazgos de nuestro estudio. - Finalmente, en la Sección 5 se extraen las principales conclusiones.

Fundamentos teóricos de la investigación

Marco conceptual

Elabora tu marco conceptual con 20 conceptos como mínimo, conceptos extraídos del desarrollo de todo el trabajo, son conceptos pequeños del diccionario, ordenados alfabéticamente.

Marco Teórico

En esta sección del trabajo, se describen y analizan los estudios llevados a cabo previamente, que tienen relación con el objeto de estudio. Algunos autores (Ibáñez Brambilla, 165) sugieren que la extensión de este apartado sea de entre 25 y 40 páginas.

Análisis bibliográfico

1) Conceptos sobre Revisión Sistemática La declaración PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses), publicada en 2009, se diseñó para ayudar a los autores de revisiones sistemáticas a documentar de manera transparente el porqué de la revisión, qué hicieron los autores y qué encontraron. Durante la última década, ha habido muchos avances en la metodología y terminología de las revisiones sistemáticas, lo que ha requerido una actualización de esta guía. La declaración PRISMA 2020 sustituye a la declaración de 2009 e incluye una nueva guía de presentación de las publicaciones que refleja los avances en los métodos para identificar, seleccionar, evaluar y sintetizar estudios. La estructura y la presentación de los ítems ha sido modificada para facilitar su implementación. En este artículo, presentamos la lista de verificación PRISMA 2020 con 27 ítems, y una lista de verificación ampliada que detalla las recomendaciones en la publicación de cada ítem, la lista de verificación del resumen estructurado PRISMA 2020 y el diagrama de flujo revisado para revisiones sistemáticas. <https://doi.org/10.1016/j.recesp.2021.06.016>

“Las revisiones sistemáticas son útiles en muchos aspectos críticos, ya que pueden proporcionar una síntesis del estado del conocimiento en un área determinada, a partir de la cual se pueden identificar futuras prioridades de investigación, abordar preguntas que de otro modo no podrían ser respondidas por estudios individuales, identificar problemas en la investigación primaria que deben ser corregidos en futuros estudios y generar o evaluar teorías sobre cómo o por qué ocurren fenómenos de interés.

Por lo tanto, las revisiones sistemáticas generan diversos tipos de conocimientos para diferentes usuarios de las revisiones (por ejemplo, pacientes, proveedores de atención médica, investigadores y responsables de la formulación de políticas de salud).

Para garantizar que una revisión sistemática sea valiosa para los usuarios, los autores deben elaborar una publicación transparente, completa y precisa en la que se describa por qué se ha realizado la revisión, qué se ha hecho (por ejemplo, cómo se han identificado y seleccionado los estudios) y qué se ha encontrado (por ejemplo, características de los estudios incluidos y los resultados de los metanálisis)

<https://doi.org/10.1016/j.recesp.2021.06.016>

” Se han producido muchas innovaciones en la realización de revisiones sistemáticas desde la publicación de la declaración PRISMA 2009. Por ejemplo, los avances tecnológicos han permitido el uso del procesamiento del lenguaje natural y el aprendizaje automático para identificar evidencia relevante (21–23), se han propuesto métodos para sintetizar y presentar los hallazgos cuando no es posible o apropiado realizar un metanálisis (24–26) y se han desarrollado nuevos métodos para evaluar el riesgo de sesgo en los resultados de los estudios incluidos (27, 289).

Se ha acumulado evidencia sobre las fuentes de sesgo en las revisiones sistemáticas, dando lugar al desarrollo de nuevas herramientas para evaluar la realización de revisiones sistemáticas (29, 30). La terminología utilizada para describir determinados procesos de las revisiones también ha evolucionado, ya que se ha pasado de evaluar la “calidad” a evaluar la “certidumbre” del cuerpo de la evidencia (31). Además, el panorama editorial se ha transformado y actualmente existen múltiples posibilidades para registrar y difundir protocolos de revisiones sistemáticas (32, 33), difundir las publicaciones de revisiones sistemáticas y compartir datos y materiales, tales como servidores de preimpresión y repositorios de acceso público. ”

<https://doi.org/10.1016/j.recesp.2021.06.016>

2) Conceptos sobre Bibliometría Clarivate Analytics (2018a) Clarivate Analytics (2018b) (Cuccurullo et al., 2016 ; Aria y Cuccurullo, 2017) (Pritchard, 1969; Broadus, 1987)

La bibliometría es un campo de investigación dentro de las ciencias de la información y bibliotecas que estudia el material bibliográfico utilizando métodos matemáticos y estadísticos (Pritchard, 1969; Broadus, 1987). A partir del crecimiento exponencial de los datos y la concetividad brindada por internet en los últimos años, la bibliometría se ha vuelto una herramienta popular para asistir en los procesos de revisión de la bibliografía e identificación de los principales tópicos, por ejemplo áreas de investigación relevantes, centros de investigación, autores, coautores y países de mayor producción, entre otros.

Un estudio bibliométrico responde al tipo observacional, enfocado en la identificación de patrones o regularidades en un corpus o grupo de producción científica, captado a partir de diferentes medios de divulgación.

Existen dos tipos de indicadores bibliométricos: “de actividad” y “de impacto”. Los indicadores de actividad, permiten evaluar: el número de publicaciones y su distribución, la productividad del autor (número de trabajos por autor, por revista o por institución), la dispersión de las publicaciones, la colaboración entre las publicaciones (número de autores por publicación, y la cooperación científica entre investigadores e instituciones), la vida media de la citación (número de años transcurridos desde la publicación en el cual las citas disminuyen a la mitad) y las conexiones entre los autores.

Mientras que los indicadores de impacto permiten determinar cuáles son los documentos más importantes en el campo de interés (documentos muy citados hasta los 2 a 4 años luego de su publicación), se mide también

el impacto de las revistas, mediante los datos divulgados por el Institute for Scientific Information (ISI) en el Journal Citation Report (JCR) (10).

Adicionalmente, dentro de los índices bibliométricos se encuentra el índice h, que conjuga, en un mismo índice bibliométrico, el parámetro de la “cantidad” (número de publicaciones) y el parámetro de la “calidad” (tasa de citas) de un modo balanceado, superando los recuentos primarios de documentos y citas. Pondera un trabajo científico prolongado durante toda la vida profesional o académica, frente a verdaderos “chispas” –éxitos fulgurantes, pero aislados– que pueden tener un impacto muy superlativo, pero claramente acotado (29). Esto permite entender que el índice h en la actualidad es criticado puesto que puede acarrear error en la interpretación debido a que los autores que recién empiezan a publicar tendrán un “bajo índice h” (por tener un número bajo de publicaciones) en contraste con aquellos que ya llevan publicando varios artículos previos, por lo que es mejor medir la cantidad de citaciones por artículo.

2.1) Herramientas de apoyo a la bibliometría: Bibliometrix, paquete de R ...

Machine learning aplicado al procesamiento de textos

1) IA Definición de IA y aplicaciones

1.1) Redes neuronales artificiales El cerebro humano es el sistema de reconocimiento de patrones más complejo y eficiente que conocemos; los humanos realizamos acciones tan sorprendentes como identificar a un conocido en una multitud o reconocer de oído el solista de un concierto para violín [1]. La inteligencia artificial intenta emular facultades intelectuales humanas a través de expresiones lógicas y esquemas abstractos [2] como las redes neuronales artificiales (RNA) cuya estructura y funcionamiento están inspirados en las redes neuronales biológicas, de ahí su nombre, construidas por un conjunto de algoritmos que intentan modelar abstracciones de alto nivel usando grafos con múltiples capas de procesamiento, dando lugar al aprendizaje automático (machine learning) y dentro de esta rama aparece el aprendizaje profundo (deep learning) [3], logrando que las unidades trabajen en conjunto, interactúen, compartan información, aprendan y apliquen el conocimiento a la resolución de nuevos casos, para lograr por ejemplo, el reconocimiento de objetos en imágenes.

El diseño de las RNA es un área en discusión, ya que no existe un modelo que resuelva todos los problemas, sino que su arquitectura define en gran parte las habilidades del sistema y su campo de aplicación, desde el modelo monocapa o perceptrón simple, formado por una capa de neuronas artificiales (nodos) que proyectan las entradas a una capa de neuronas de salida, hasta modelos híbridos que surgen de la combinación de distintos tipos de redes, como por ejemplo el modelo convolucional (CNN) donde las neuronas de una capa solamente se unen con un subgrupo de ellas, buscando reducir el número de neuronas y el costo computacional y el modelo recurrente (RNN) que permite conexiones arbitrarias entre las neuronas, incluso en forma de ciclos, logrando simular condiciones de temporalidad y memoria [4].

1.2) Redes neuronales artificiales y reconocimiento de patrones y secuencias Las CNN trabajan dividiendo y modelando la información en partes más pequeñas, y combinando esta información en las capas más profundas de la red. Por ejemplo, en el caso del tratamiento de una imagen, las primeras capas tratarían de detectar los bordes de las figuras, las siguientes capas buscarían combinar los patrones de detección de bordes para conseguir formas más simples y además se puede sumar conocimiento a partir de otros patrones como la posición de los objetos o la iluminación. Finalmente en las últimas capas se podría hacer coincidir la imagen con todos los patrones descubiertos, para conseguir una predicción final considerando la suma de todos ellos. Así es como las redes neuronales convolucionales consiguen modelar una gran cantidad de datos, dividiendo previamente el problema en partes para conseguir predicciones más sencillas y precisas [5].

Las RNN por su parte aplican mejor a aquellos casos donde la información reside en el análisis temporal, contando con datos en la forma de series o secuencias, de esta forma los estados de la red en un instante

temporal se nutren de información de los estados anteriores, destacando el buen desempeño de las RNN-LSTM (Long Short Term Memory) y GRU (Gated Recurrent Units) [6] en los casos donde se desea incorporar información de las dependencias a largo plazo.

1.3) Redes híbridas como opción para mejorar la capacidades de los modelos puros Las áreas de aplicación donde se requiere resolver problemas de clasificación y predicción [7] cuya solución involucra la detección de patrones y codificación de secuencias [8] parecen hechos a la medida de la IA, obteniendo logros en campos como procesamiento de imágenes y lenguaje natural [9], traducción automática [10], reconocimiento de rostros [11] y realidad virtual [12].

En este contexto surgen tendencias como la denominada industria 4.0 a través de la integración de técnicas como machine learning, big data, internet de las cosas con el objetivo de aportar inteligencia a los procesos [13], por ejemplo en las empresas de producción de bienes la identificación automática de objetos y la predicción de acciones pueden conducir a la mejora sustancial de la productividad.

Impulsadas por las potencialidades, surgen diversas líneas de investigación en relación a las RNN, algunas se centran en ajustar los parámetros de una topología pura (capas, entradas, salidas) [14], por ejemplo las CNN muestran buenos resultados para resolver problemas relacionados con el tratamiento de imágenes, como clasificación, reconocimiento, recuperación, segmentación y detección de objetos [15] pero sin resultados significativos en la identificación de las acciones que realizan estos objetos.

Por su parte las RNN están diseñadas para captar información temporal, mostrando resultados prometedores en la codificación de información a partir de una secuencia de imágenes [16], otras líneas proponen generar un nuevo modelo a partir de la combinación de modelos puros, con el objeto de aprovechar las capacidades de cada tipo, por ejemplo el caso de la combinación de redes convolucionales (para la detección de objetos) y recurrentes (para la codificación de acciones) [17], [18].

2) Machine learning Concepto de Machine Learning y aplicaciones

2.1) Machine Learning y procesamiento de textos Concepto de Machine Learning y procesamiento de textos

Hipótesis y Variables

Formulación de hipótesis

Identificación de Variables

Clasificación de variables

Definición constitutiva de variables

Definición operacional de variables

Descripción de las variables.

Estructura de los datos

Describir brevemente los datos (tidy data): - Cuántas variables tienen? - De qué tipo es cada variable? - Cuántas observaciones?

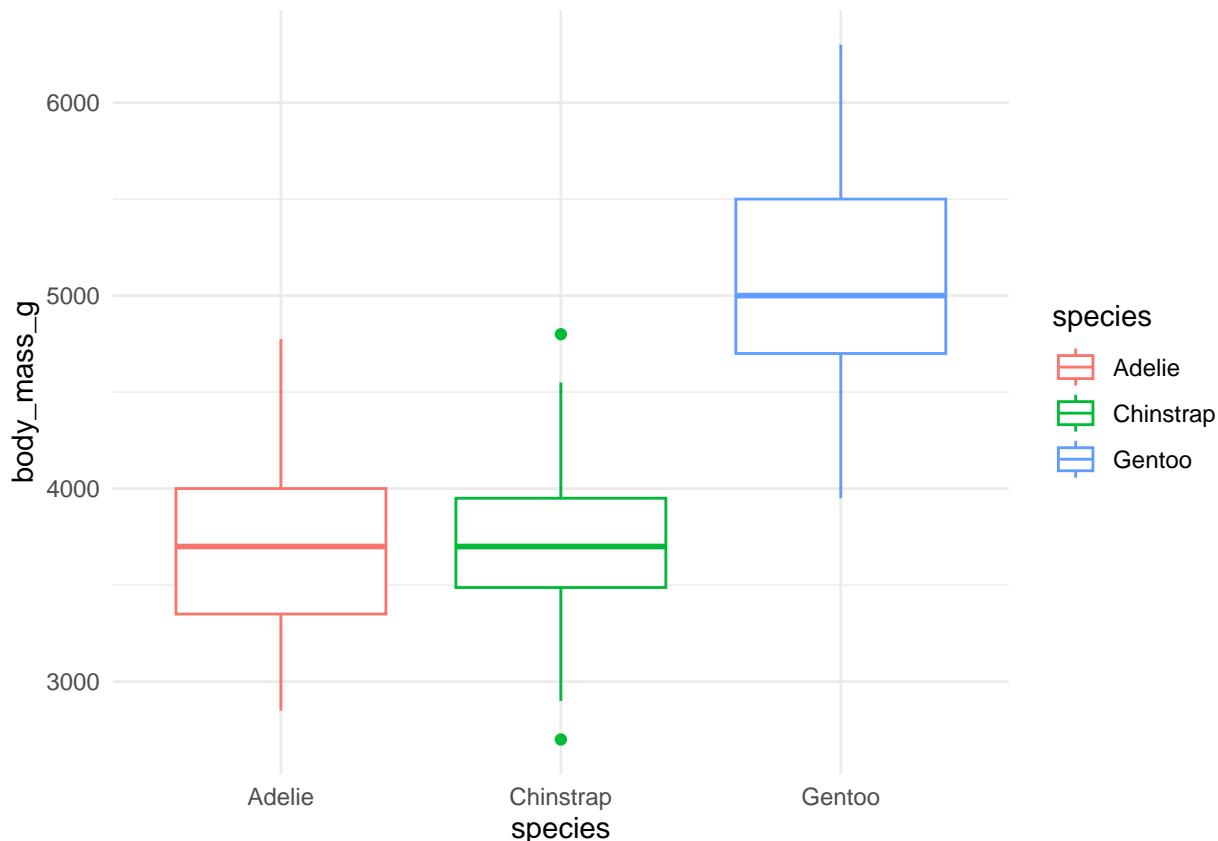
En caso de que ya tengan los datos en formato tidy como para cargar en **R**, se puede acompañar con algún breve análisis exploratorio. Por ejemplo:

```
library(palmerpenguins)
str(penguins)

## # tibble [344 x 8] (S3:tbl_df/tbl/data.frame)
## $ species      : Factor w/ 3 levels "Adelie","Chinstrap",...
## $ island       : Factor w/ 3 levels "Biscoe","Dream",...
## $ bill_length_mm: num [1:344] 39.1 39.5 40.3 NA 36.7 ...
## $ bill_depth_mm: num [1:344] 18.7 17.4 18 NA 19.3 20.6 ...
## $ flipper_length_mm: int [1:344] 181 186 195 NA 193 190 ...
## $ body_mass_g   : int [1:344] 3750 3800 3250 NA 3450 3650 ...
## $ sex          : Factor w/ 2 levels "female","male": 2 1 1 ...
## $ year         : int [1:344] 2007 2007 2007 2007 2007 2007 ...

penguins %>% ggplot(aes(x = species,
                           y = body_mass_g,
                           color = species)) +
  geom_boxplot() +
  theme_minimal()

## Warning: Removed 2 rows containing non-finite values ('stat_boxplot()').
```



Operacionalización de variables

Variable Tipo Opción Año de publicación Numérica discreta No aplica

Nombre de revista Categórica nominal No aplica Cuartil de publicación de revista según Scimago, en el año de publicación Categórica ordinal 0 = No indexado en Scopus 1 = Q1 2 = Q2 3 = Q3 4 = Q4 Idioma de artículo Categórica nominal 0 = Inglés 1 = Español 2 = Español e inglés 3 = Portugués 4 = Otro País de filiación institucional de autores Categórica nominal No aplica Autor corresponsal Categórica nominal No aplica Filiación institucional del autor corresponsal Categórica nominal No aplica Tipo de estudio Categórica nominal 0 = Transversal 1 = Cohorte o caso control 2 = Cuasi-experimental 3 = Ensayo Clínico 4 = Revisión sistemática 5 = Otros Especialidad Categórica nominal No aplica Temática Categórica nominal No aplica Área médica (Área sanitaria de aplicativo) Categórica nominal 0 = Preventivo-promocional 1 = Diagnóstico 2 = Terapéutico-Control 3 = Rehabilitación 4 = Gestión en salud o salud pública 5 = Educación médica Población objetivo Categórica nominal 0 = Paciente 1 = Estudiantes de salud 2 = Personal de salud médico 3 = Personal de salud no médico Financiamiento del estudio Categórica nominal 0 = No declarar 1 = Autofinanciado 2 = Público (Estatal) 3 = Privado (Empresas) 4 = Entidad educativa y/o investigación

Colaboración entre países de Latinoamérica Categórica nominal 0 = No colaboración 1 = Existe colaboración (si existe un artículo como mínimo cuyas filiaciones de los autores pertenezcan a más de un país latinoamericano)

Colaboración entre países de Latinoamérica con otros países Categórica nominal 0 = No colaboración 1 = Existe colaboración (si existe cinco (05) artículos como mínimo cuyas filiaciones de los autores pertenezcan a más de un país latinoamericano)

Metodología

En esta parte de la tesis se reportan y se justifican los métodos, participantes e instrumentos utilizados para la investigación. La descripción de la metodología debe ser breve y clara, ya que se busca que otro investigador pueda repetir el estudio sin necesidad de cuestionar al autor. En esta parte del proyecto, también suelen describirse los métodos que se usarán en el análisis de datos.

Este capítulo comprende los siguientes apartados: sujetos, material y procedimiento. Los sujetos son las personas sobre las que se realizó el estudio. En la sección de metodología se describen las características de la población seleccionada, y se menciona la razón por la cual fueron elegidos. El material que se utilizó para el proyecto también debe ser reportado; si se usaron cuestionarios, o se hicieron entrevistas o se usó cualquier otro instrumento de investigación también ha de especificarse. Finalmente, es necesario detallar el procedimiento que se llevó a cabo para realizar el trabajo.

Diseño y Tipo del Estudio

Estudio de análisis bibliométrico. La metodología aplicada fue de tipo descriptivo con enfoque cuantitativo. Se usaron indicadores bibliométricos para realizar el estudio. El primer paso fue desarrollar la ecuación de búsqueda y probarla en la base de datos utilizada, que fue SCOPUS; luego de obtener los resultados se aplicaron los indicadores bibliométricos; después, los resultados se analizaron y se obtuvieron las conclusiones.

Población

Artículos científicos sobre aplicativos móviles en salud en Latinoamérica publicados en la base de datos de Scopus y LILACS. Se consideró Scopus debido a que alberga revistas científicas de otras ciencias aparte de la médica, que también podrían incluir artículos científicos que evalúen aplicativos móviles en salud. Adicionalmente LILACS es uno de los registros más importantes que incluye producción científica sólo de Latinoamérica.

Muestra

La muestra está definida por la totalidad de la población luego de aplicación de los criterios de inclusión y exclusión.

Flujograma de la Metodología

(A continuación un esquema del diagrama de flujo)

Revisar y adecuar los pasos anteriormente planteados

- Identificación Registros incluidos usando términos de búsqueda relacionados con “IA DL Video” en Base de datos IEEE ExPlore entre los años xxxx-2022 (n= xxx.xxx)

Registros excluidos (n = xxx.xxxx) xxx.xxx no incluyeron país xxx registros publicados fuera del periodo de análisis xxx Registros duplicados

Primer separación (por ejemplo por tema)

- Screening (artículos que quedan luego del filtro) Artículos analizados (n=xxxx)

Segunda separación (por subtema)

- Included (clasificación por algún tipo) Artículos con investigaciones originales (n=xxxx) Revisiones narrativas (n=xxxx) Cartas de editorial (n=xxx) Otros (n=xxx)

Criterios de inclusión y exclusión

Criterios de inclusión

- Artículos originales que se encuentren en formato IMRD (Introducción, Métodos, Resultados y Discusión), o sus similares, incluyendo revisiones sistemáticas, publicadas en revistas científicas.
- Artículos que evalúen la efectividad, describan el desarrollo o describan las características de al menos un programa tipo aplicativo disponible en un aparato móvil (aplicativo móvil) cuyo objetivo original esté relacionado con la salud, independientemente del tipo de usuario que la aplica (Preventivo-promocional, diagnóstico, tratamiento, recuperación, o administración en salud).
- Artículos que tengan al menos un autor con filiación institucional de algún país catalogado como Latinoamérica según el Scimago Country Ranking que incluye a Antigua y Barbuda, Antillas Holandesas, Argentina, Aruba, Bahamas, Barbados, Belice, Bermudas, Bolivia, Brasil, Colombia, Costa Rica, Chile, Cuba, Dominica, Ecuador, El Salvador, Granada, Guadalupe, Guatemala, Guyana, Guyana Francesa, Haití, Honduras, Islas Caimán, Islas Malvinas, Islas Vírgenes, Jamaica, Martinica, México, Montserrat, Nicaragua, Panamá, Paraguay, Perú, Puerto Rico, República Dominicana, San Cristóbal y Nieves, Santa Lucía, Surinam, Trinidad y Tobago, Uruguay y Venezuela (34).
- Artículos publicados hasta el día en el que se realizó la búsqueda sistemática: 27 de diciembre 2019.

Criterios de exclusión

- Artículos tipo editoriales, resúmenes de congresos, “proceeding papers”, o cartas al editor.
- Artículos con información secundaria tipo revisiones narrativas.
- Tesis de grado o documentos técnicos
- Artículos no disponibles a texto completo

Técnicas y procedimiento

A continuación se describen el contexto en el cual se desarrolla el estudio bibliométrico.

1) Descripción del Laboratorio

- 1) Describir la técnica de extracción a utilizar. En principio será la consulta a BD a través de APIs, para facilitar la reproducibilidad del experimento. Por lo cual se reduce el universo de candidatos a aquellos que dispongan de APIs abiertas.
- 2) Descripción del hardware y software a utilizar

2) Selección de las BD a consultar

- 1) Definición de la/las bases de datos. En este caso WebofScience y otras que dispongan de APIs abiertas.
- 2) Descripción de los filtros configurables en la API disponible para cada BD (armar una tabla comparativa).
- 3) Análisis de las características del set de datos obtenido (nombres y formato de los metadatos, entre otros). A fin de apoyar el proceso de homogenización de los distintos set de datos.

3) Definición y análisis de los metadatos de interés.

- 1) Definición del conjunto coincidencia entre los metadatos deseables y los filtros disponibles en las APIs.
- 2) Estandarización de campos (para facilitar el proceso de unificación de los sets de datos obtenidos por las consultas a las distintas BD definidas).
- 3) Estandarización del formato de salida (se define el estándar csv)

4) Descripción de los criterios de extracción, transformación y carga (ETL)

ETL es un tipo de integración de datos que se refiere a los tres pasos (extracción, transformación, carga) utilizados para combinar datos de múltiples fuentes. A menudo se usa para construir un almacén de datos. Durante este proceso, los datos se toman (extraen) de un sistema de origen, se convierten (transforman) en un formato que se puede analizar y almacenar (cargar) en un almacén de datos u otro sistema. Extraer, cargar, transformar (ELT) es un enfoque alternativo pero relacionado diseñado para empujar el procesamiento a la base de datos para mejorar el rendimiento.

Extracción La primera parte del proceso ETL consiste en extraer los datos desde los sistemas de origen. La mayoría de los proyectos de almacenamiento de datos funcionan con datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o ficheros planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

Definición de la Estrategia de búsqueda Se diseñó una estrategia de búsqueda para cada base de datos de manera individual, con el objetivo de recolectar todos los artículos que tengan los términos de “aplicativos móviles” o sus similares en el título, resumen o cómo palabra clave.

Construcción de la ecuación de búsqueda. Para la construcción de la ecuación de búsqueda se tomaron de base los temas de la investigación, que son: Inteligencia Artificial, Política Pública, Toma de Decisiones y Agricultura. A partir de este punto se buscaron las palabras clave que los autores y las publicaciones utilizaban para referenciar el tema. Al final se obtuvo la ecuación, con la cual se obtuvieron 110 publicaciones. La ecuación de búsqueda se presenta a continuación.

Configurar las APIs según los criterios definidos en la Etapa 3

Ejecutar el proceso de extracción

Transformación La fase de transformación aplica una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. ##### Validar metadatos perdidos ##### Validar que no existan elementos repetidos Por ejemplo un autor que aparece con distintas versiones del nombre, nombre largo y corto. ##### Validar homogenidad de las variables en los sets de datos provenientes de las distintas fuentes Se deben analizar las variables de los distintos set de datos buscando una estructura homogéna.

Carga La fase de carga es el momento en el cual los datos de la fase anterior (transformación) son cargados en el sistema de destino.

5) Validación de performance

Se debe definir el tiempo de procesamiento admisible y los criterios de particionado de los conjuntos de datos. Se debe definir la variable que se utilizará para el particionado, en caso que estén disponibles variables temporales, se recomienda considerar la partición centrados en escala temporal. Durante la etapa, se valida que el conjunto de datos cargado cumpla con los criterios de ejecución establecidos. En caso que no se cumpla se deberá realizar sucesivos particionados a la mitad, hasta lograr que el tiempo de procesamiento se encuentre dentro del límite.

6) Flujo de tamizaje de artículos

Título y resumen ¿Artículo original? Si -> Incluido No -> Excluido

¿Objetivo relacionado con la App? Si -> Incluido No -> Excluido

¿App con objetivo relacionado a Salud? Si -> Incluido No -> Excluido

Img 1. Flujo de tamizaje de artículos

Análisis de resultados

Una vez que se tuvieron todos los artículos incluidos, se procedió a la extracción de variables, utilizando bibliometrix. Adicionalmente, la base de datos resultante se exportó al programa de visualización VOSviewer v1.6.519 (programa libre) (40). Se obtuvieron gráficos con los conglomerados de colaboración entre países, donde el tamaño de los círculos representará la cantidad de artículos y el grosor de las líneas indica el número de artículos en colaboración. Adicionalmente, se utilizó el programa QGIS v2.0 (programa libre) para graficar la distribución geográfica de artículos por país en Latinoamérica.

Aplicación de indicadores bibliométricos: Con los datos bibliográficos de las publicaciones que resultaron, se aplicarán los siguientes indicadores bibliométricos.

En este apartado se plantean los resultados obtenidos, acompañados de su respectiva explicación y un análisis profundo. Se pueden utilizar tablas y gráficas para reportar los resultados, si esto facilita su comprensión. El objetivo de esta sección es mostrar si los datos obtenidos apoyan la hipótesis de la investigación (Ibáñez Brambilla, 171). Las tablas y gráficas que se utilicen deberán ser tituladas y numeradas según su orden de aparición. Los títulos han de ser claros, precisos y no repetitivos.

Presentar lo encontrado: patrones o regularidades en un corpus o grupo de producción científica grande (esto puede ser red de coautores, importancia de determinados países en un determinado ámbito, entre otros)

Se obtuvieron 844 citaciones a partir de las dos búsquedas iniciales en Scopus y LILACS. Se excluyeron 33 citaciones obtenidas por la búsqueda en LILACS debido a que eran repetidas o eran tesis de grado. Se tamizaron 811 artículos para evaluar si cumplían con los criterios de selección. Se excluyeron 540 artículos, obteniendo finalmente 271 artículos incluidos en el presente análisis (Figura 2)

PRISMA 2020, propone evaluar Estudios incluidos - Proporcionar el número total de estudios incluidos. - Proporcionar el número total de participantes. - Resumir las características relevantes de los estudios.

Los resultados obtenidos con los indicadores bibliométricos se analizaron y se obtuvieron las conclusiones. Apoyados por herramientas de visualización, por ejemplo VOSviewer.

El presente trabajo mostró el aumento progresivo de publicaciones científicas sobre aplicativos móviles en salud en Latinoamérica desde el 2012. Más de la mitad de los artículos analizados tienen al menos un autor de Brasil. De manera similar, la mayoría de los autores correspondientes identificados eran provenientes de universidades principalmente de Brasil. Finalmente, se evidencia una colaboración en investigación de Brasil con México, y de Brasil con EUA. Hasta nuestro conocimiento, este es el primer análisis bibliométrico de la actividad de investigación publicado en revistas indexadas sobre aplicativos móviles en salud.

Según PRISMA 2020: En cuanto a la discusión propuesta, según PRISMA, se recomienda destacar las “Limitaciones de la evidencia” (Proporcione un breve resumen de las limitaciones de la evidencia incluida en la revisión (por ejemplo, riesgo de sesgo, inconsistencia –heterogeneidad– e imprecisión) y además destacar en la Interpretación (Proporcione una interpretación general de los resultados y sus implicaciones importantes).

Este documento trabaja con datos de Web of Science Core Collection (WoS), Clarivate Analytics. Seleccionamos todos los trabajos indexados que contienen ‘bitcoin’ como tema, lo que hace un total de 1162 documentos, publicados en 703 fuentes (revistas, libros, etc.), durante el período 2012-2019. Estos documentos fueron (co-)escritos por 2293 personas. La gran mayoría de los documentos son de varios autores, siendo sólo 322 documentos de un solo autor. El número medio de autores por documento es de 1,97.

Bitcoin como tema de investigación comprende varias disciplinas.

Web of Science asigna artículos indexados a una o más áreas de investigación. Los 1162 artículos considerados en nuestro muestra se asignaron 1543 áreas de investigación. Las cinco principales áreas de investigación se muestran en la Tabla 1 Principales áreas de investigación asignadas a los trabajos de la muestra (Áreas de investigación, Registros, % de 1543).

Tabla 2 Número de artículos publicados por año, con ‘bitcoin’ como tema. Fuente: Web of Science Colección Núcleo (Año, # artículos, Tasa de crecimiento anual)

Indicadores bibliométricos

Con los datos bibliográficos de las publicaciones que resultaron, se aplicarán los siguientes indicadores bibliométricos.

1) Sobre las Áreas de Investigación

Identificar las Áreas de investigación estudiadas. A fin de enfocarse en las áreas que producen la mayor cantidad de artículos. (Tabla 1)

Tabla 1. Principales áreas de investigación asignadas en los trabajos en el ejemplo (Areas de investigación, # Registros, % del total)

2) Sobre los Años de publicación

Histórico de publicaciones. Este indicador muestra el número de publicaciones realizadas año a año.

Identificar los años de explosión de artículos (Yearly growth rate)

3) Sobre el país de origen de la producción

Distribución geográfica del autor Tabla 3 Los diez principales países de los autores correspondientes (País, Artículos, Frecuencia, Publicaciones de un solo país, Publicaciones de múltiples países, MCP_Ratio)

País de origen. Este presenta el número de publicaciones por país, acorde al país donde se realizó la publicación.

4) Sobre las Citas

Distribución de Citas totales por país Tabla 4 Top ten de citas totales por país (País, Citas totales, Promedio de citas de artículos).

Citación de autores. Este indicador presenta cómo se relacionan y son citados los autores. El método utilizado para encontrar las relaciones y los grupos es el ya nombrado fuerza de asociación, para lo cual se usó también el software VosViewer.

Identificación de la colaboración internacional Además, alrededor del 35% de los artículos publicados son de autores de múltiples países. La colaboración internacional en EE. UU. e India están por debajo del promedio, con una tasa de publicaciones en varios países del 14 % y el 10 %, respectivamente. El internacional la colaboración alcanza un máximo para España, donde el 46% de los trabajos son de este tipo.

5) Sobre la fuente de publicación

Tipo de publicación. Este indicador muestra el número de publicaciones clasificadas por tipo de medio, por ejemplo revista especializada, memoria de conferencia, revisión, libro o capítulo de libro.

Identificar las revistas y libros más importantes sobre el tema. Es importante identificar las fuentes de producción cuando se está haciendo un artículo bibliométrico, porque sería un potencial espacio donde publicar el artículo que se está realizando)

Principales fuentes de publicación Tabla 5 Las diez fuentes más relevantes (Fuentes #Artículos Tipo).

Incluir el análisis de las revistas. Tabla 6 (Revista, % Documentos publicados, # Citas, Citas/Documentos, cite Score, Quartil, SIR, % Colaboracion internacional)

6) Sobre los artículos

Artículos muy citados Tabla 7 Artículos altamente citados, en orden descendente por número de citas (Autor (año), Título, Fuente, # Citas).

7) Sobre las Palabras Clave Para las palabras clave se usó el método de fuerza de asociación. Este método busca determinar cuáles categorías se relacionan más entre ellas y se agrupan acorde a ello. Para ello se usó el software libre VosViewer, el cual es un software para análisis bibliométrico que ayuda a crear redes bibliométricas para ver gráficamente estas relaciones.

Palabras clave principales. Tabla 6 Palabras clave principales (Palabras clave del autor (DE), #Artículos, Palabras clave+ (ID), #Artículos)

8) Sobre los autores

Detectar los autores más productivos (con mayor cantidad de artículos) Este muestra el listado de autores y el número de publicaciones científicas de cada uno. Tabla 8 Autores más productivos (Autores, Institución, # artículos).

Índice de concentración de autores, como una especie de entropía, definido en el artículo publicado en IDEAS.

Grado de concentración de las variables seleccionadas Propuesta en el artículo ...

En este subapartado analizamos globalmente varias variables bibliométricas, con el fin de mostrar el grado de concentración de las mismas. Una característica importante en los estudios bibliométricos es la uniformidad de la contribución de los autores, países y revistas a una investigación. La teoría de la información proporciona métricas alternativas a las medidas estadísticas tradicionales de concentración, como la desviación estándar, asimetría o curtosis. En particular, Shannon y Weaver (1949) desarrollaron la célebre entropía de Shannon. Dada una distribución de probabilidad de tributación

Tabla 9 Índice de concentración entrópica () de seleccionados variables (Variable, H)

Tabla 10 Distribución observada del número de autores que escribieron un número determinado de artículos y valores ajustados por la ley de Lotka (# Artículos, # Autores, Frecuencia observada, Frecuencia ajustada)

9) Sobre las organizaciones a las que pertenecen los autores

Afiliación de autores. Este muestra las organizaciones a las cuales pertenecen los autores y el número de publicaciones asociadas a cada organización.

10) Sobre los patrocinadores

Patrocinador Este indica los patrocinadores que financiaron las investigaciones que generaron las publicaciones. El número corresponde a las publicaciones que tiene asociadas cada patrocinador.

Análisis de las visualizaciones

- 1) Enfocados en los artículos. Representando los artículos por nodos, junto con el tamaño del nodo representando la cantidad de citas que recibe ese nodo. El nodo se etiqueta con la palabra clave más característica. Análisis de las visualizaciones (por ejemplo, en el artículo de bitcoin se observó que hay tres cluster importantes, uno más vinculado a business economics, otra que era mas computer science pero divididos en una parte que se dedicaba a blockchain duro (más teórico) y una parte más aplicada a los protocolos que se aplicaban a criptomonedas en concreto y bitcoin en particular).
- 2) Enfocados en los autores. Donde se destaque los autores y coautores mediante citas.

Gráficos de citas, fuentes y autores Las siguientes cifras fueron generadas utilizando el software VOS viewer, que permite contar las palabras que aparecen en el título, resumen y palabras clave para construir todas las relaciones que aparecen entre los diferentes documentos publicados en la Web of Science (van Eck y Waltman, 2010).

Fig. 1. Mapa de nubes de palabras en títulos y resúmenes (conteo completo), generado con VOSviewer (<http://www.vosviewer.com/>). (Para interpretación de las referencias al color en el texto, se remite al lector a la versión web de este artículo).

Fig. 2. Mapa de nubes de palabras en títulos y resúmenes (conteo binario), generado con VOSviewer (<http://www.vosviewer.com/>). (Para interpretación de las referencias al color en el texto, se remite al lector a la versión web de este artículo).

Fig. 3. Mapa en la nube de revistas donde se publicaron artículos sobre bitcoin, generado con VOSviewer (<http://www.vosviewer.com/>).

Fig. 4. Mapa en la nube de revistas de autores con artículos sobre bitcoin publicados, generado con VOSviewer (<http://www.vosviewer.com/>).

Resultados

Síntesis de los resultados

- Presentar los resultados de los desenlaces principales
- Indicar, preferiblemente, el número de estudios incluidos y los participantes en cada uno de ellos.
- Si se ha realizado un metanálisis, indicar el estimador de resumen y el intervalo de confianza o de credibilidad.
- Si se comparan grupos, describa la dirección del efecto (por ejemplo, qué grupo se ha visto favorecido).

Tabla típica de resultados. Año de publicación y tipos de artículos (Año, # artículos, Investigaciones originales, Artículos de revisión, Cartas, Otros, p value) Incorporar rango de años, por ejemplo 5 años, y totalizar al final Otros: conference paper, note, editorial, short survey, erratum.

Plan de análisis

Una vez controlada la calidad de los datos se procedió a realizar un análisis descriptivo de nuestros resultados. Las variables fueron resumidas según su frecuencia relativa y frecuencia absoluta. La producción científica de cada país latinoamericano fue usada para calcular el índice de producción científica según el tamaño de población durante el 2019 (Disponible en: <https://population.un.org/wpp/DataQuery/>) y según el promedio de investigadores dedicados a investigación y desarrollo por cada millón de personas por país durante 2012 a 2017 (Disponible en: <https://datos.bancomundial.org/indicador/SP.POP.SCIE.RD.P6>). El análisis de los datos se realizó utilizando el paquete estadístico STATA MP v16 (Statacorp, Texas, USA)

Adicionalmente, la base de datos resultante se exportó al programa de visualización VOSviewer v1.6.519 (programa libre) (40). Se obtuvieron gráficos con los conglomerados de colaboración entre países, donde el tamaño de los círculos representará la cantidad de artículos y el grosor de las líneas indica el número de artículos en colaboración. Adicionalmente, se utilizó el programa QGIS v2.0 (programa libre) para graficar la distribución geográfica de artículos por país en Latinoamérica.

1) Descripción del Laboratorio

- 1) Describir la técnica de extracción a utilizar. En principio será la consulta a BD a través de APIs, para facilitar la reproducibilidad del experimento. Por lo cual se reduce el universo de candidatos a aquellos que dispongan de APIs abiertas.

2) Descripción del hardware y software a utilizar

2) Selección de las BD a consultar

Se realizó una búsqueda sistemática en las bases de datos Scopus y Literatura Latinoamericana y del Caribe en Ciencias de la Salud (LILACS), Scopus (<https://www.scopus.com/>) incluye más de 22000 revistas de ciencia, tecnología, ciencias sociales, artes, humanidades y medicina.

Previamente, algunos estudios de tipo bibliométrico (11, 35, 36), utilizaron solamente la base de datos Scopus, por ser una base que tiene un motor de búsqueda con herramientas para obtener los artículos científicos según filiación institucional, país y nombre de los autores, así como según la revista científica y tipo de artículo, asimismo incluye a todos los artículos que se acceden a través de la base de datos Medline (37, 38).

Adicionalmente, LILACS (<https://lilacs.bvsalud.org/es/>) es la base de datos más importante en Latinoamérica y el Caribe. Es administrada por la Organización Panamericana de Salud, e incluye más de 880.000 registros de artículos de revistas científicas, tesis, documentos gubernamentales, resúmenes de congresos y libros, desde 1982 (39).

Se consideraron todos los artículos que tengan los términos “aplicativos móviles” o sus similares en el título, resumen o como palabra clave y que tengan por lo menos un autor con filiación de un país latinoamericano.

La búsqueda se realizó el 27 de diciembre del 2019 en ambas bases de datos. Luego, se realizó la exclusión de los documentos duplicados de manera manual. Posterior a ello, dos investigadores independientes realizaron la revisión del título y resumen de los artículos, la calificación de estos investigadores usó como criterio el flujo de tamizaje (Figura 1), con lo que se determinó la inclusión de los artículos evaluados.

3) Definición y análisis de los metadatos de interés.

- 1) Definición del conjunto coincidencia entre los metadatos deseables y los filtros disponibles en las APIs.
- 2) Estandarización de campos (para facilitar el proceso de unificación de los sets de datos obtenidos por las consultas a las distintas BD definidas).
- 3) Estandarización del formato de salida (se define el estándar csv)

4) Descripción de los criterios de extracción, transformación y carga (ETL)

ETL es un tipo de integración de datos que se refiere a los tres pasos (extracción, transformación, carga) utilizados para combinar datos de múltiples fuentes. A menudo se usa para construir un almacén de datos. Durante este proceso, los datos se toman (extraen) de un sistema de origen, se convierten (transforman) en un formato que se puede analizar y almacenar (cargar) en un almacén de datos u otro sistema. Extraer, cargar, transformar (ELT) es un enfoque alternativo pero relacionado diseñado para empujar el procesamiento a la base de datos para mejorar el rendimiento.

Extracción La primera parte del proceso ETL consiste en extraer los datos desde los sistemas de origen. La mayoría de los proyectos de almacenamiento de datos funcionan con datos provenientes de diferentes sistemas de origen. Cada sistema separado puede usar una organización diferente de los datos o formatos distintos. Los formatos de las fuentes normalmente se encuentran en bases de datos relacionales o ficheros planos, pero pueden incluir bases de datos no relacionales u otras estructuras diferentes. La extracción convierte los datos a un formato preparado para iniciar el proceso de transformación.

Definición de la Estrategia de búsqueda Se diseñó una estrategia de búsqueda para cada base de datos de manera individual, con el objetivo de recolectar todos los artículos que tengan los términos de “aplicativos móviles” o sus similares en el título, resumen o como palabra clave.

Construcción de la ecuación de búsqueda. Para la construcción de la ecuación de búsqueda se tomaron de base los temas de la investigación, que son: Inteligencia Artificial, Política Pública, Toma de Decisiones y Agricultura. A partir de este punto se buscaron las palabras clave que los autores y las publicaciones utilizaban para referenciar el tema. Al final se obtuvo la ecuación, con la cual se obtuvieron 110 publicaciones. La ecuación de búsqueda se presenta a continuación.

Estrategia de búsqueda para LILACS

tw:(mobile app OR smartphone app OR health app OR cellphone app) AND (db:(“LILACS”))

Estrategia de búsqueda para Scopus

TITLE-ABS-KEY (“Mobile App” OR ”Portable Electronic App” OR “Portable Software App” OR ”smart-phone app” OR ”app-based intervention” OR “mhealth app” OR ”Mobile Health App” OR “Health app” OR “Health apps” OR “phone* app” OR “cellphone app” OR “Cellular phone app*”) AND (AFFILCOUNTRY (argentina) OR AFFILCOUNTRY (bolivia) OR AFFILCOUNTRY (brazil) OR AFFILCOUNTRY (colombia) OR AFFILCOUNTRY (chile) OR AFFILCOUNTRY (ecuador) OR AFFILCOUNTRY (guyana) OR AFFILCOUNTRY (“french Guiana”) OR AFFILCOUNTRY (paraguay) OR AFFILCOUNTRY (peru) OR AFFILCOUNTRY (suriname) OR AFFILCOUNTRY (uruguay) OR AFFILCOUNTRY (venezuela) OR AFFILCOUNTRY (belize) OR AFFILCOUNTRY (“costa rica”) OR AFFILCOUNTRY (“el Salvador”) OR AFFILCOUNTRY (guatemala) OR AFFILCOUNTRY (honduras) OR AFFILCOUNTRY (nicaragua) OR AFFILCOUNTRY (panama) OR AFFILCOUNTRY (mexico) OR AFFILCOUNTRY (cuba) OR AFFILCOUNTRY (“dominican republic”) OR AFFILCOUNTRY (haiti) OR AFFILCOUNTRY (jamaica) OR AFFILCOUNTRY (“Puerto rico”) OR AFFILCOUNTRY (“trinidad and tobago”) OR AFFILCOUNTRY (barbados) OR AFFILCOUNTRY (guadeloupe) OR AFFILCOUNTRY (grenada) OR AFFILCOUNTRY (martinique) OR AFFILCOUNTRY (bermuda) OR AFFILCOUNTRY (bahamas) OR AFFILCOUNTRY (“saint kitts and nevis”) OR AFFILCOUNTRY (“netherlands antilles”) OR AFFILCOUNTRY (“falkland islands”) OR AFFILCOUNTRY (dominica) OR AFFILCOUNTRY (“cayman islands”) OR AFFILCOUNTRY (“virgin islands”) OR AFFILCOUNTRY (“Antigua and barbuda”) OR AFFILCOUNTRY (“saint lucia”) OR AFFILCOUNTRY (aruba) OR AFFILCOUNTRY (montserrat OR “Saint Vincent and the Grenadines” OR “Anguilla” OR “Turks and Caicos Islands” OR “South Georgia and the South Sandwich Islands”)) AND (EXCLUDE (DOCTYPE , “cp”)) AND (EXCLUDE (PUBYEAR , 2020))

Configurar las APIs según los criterios definidos en la Etapa 3

Ejecutar el proceso de extracción

Transformación La fase de transformación aplica una serie de reglas de negocio o funciones sobre los datos extraídos para convertirlos en datos que serán cargados. ##### Validar metadatos perdidos ##### Validar que no existan elementos repetidos Por ejemplo un autor que aparece con distintas versiones del nombre, nombre largo y corto. ##### Validar homogenidad de las variables en los sets de datos provenientes de las distintas fuentes Se deben analizar las variables de los distintos set de datos buscando una estructura homogéna.

Carga La fase de carga es el momento en el cual los datos de la fase anterior (transformación) son cargados en el sistema de destino.

5) Validación de performance

Se debe definir el tiempo de procesamiento admisible y los criterios de particionado de los conjuntos de datos. Se debe definir la variable que se utilizará para el particionado, en caso que estén disponibles variables temporales, se recomienda considerar la partición centrados en escala temporal. Durante la etapa, se valida

que el conjunto de datos cargado cumpla con los criterios de ejecución establecidos. En caso que no se cumpla se deberá realizar sucesivos particionados a la mitad, hasta lograr que el tiempo de procesamiento se encuentre dentro del límite.

6) Flujograma de tamizaje de artículos

Título y resumen ¿Artículo original? Si -> Incluido No -> Excluido
¿Objetivo relacionado con la App? Si -> Incluido No -> Excluido
¿App con objetivo relacionado a Salud? Si -> Incluido No -> Excluido

Img 1. Flujograma de tamizaje de artículos

Análisis de resultados

Una vez que se tuvieron todos los artículos incluidos, se procedió a la extracción de variables, utilizando bibliometrix. Adicionalmente, la base de datos resultante se exportó al programa de visualización VOSviewer v1.6.519 (programa libre) (40). Se obtuvieron gráficos con los conglomerados de colaboración entre países, donde el tamaño de los círculos representará la cantidad de artículos y el grosor de las líneas indica el número de artículos en colaboración. Adicionalmente, se utilizó el programa QGIS v2.0 (programa libre) para graficar la distribución geográfica de artículos por país en Latinoamérica.

Aplicación de indicadores bibliométricos: Con los datos bibliográficos de las publicaciones que resultaron, se aplicarán los siguientes indicadores bibliométricos.

En este apartado se plantean los resultados obtenidos, acompañados de su respectiva explicación y un análisis profundo. Se pueden utilizar tablas y gráficas para reportar los resultados, si esto facilita su comprensión. El objetivo de esta sección es mostrar si los datos obtenidos apoyan la hipótesis de la investigación (Ibáñez Brambilla, 171). Las tablas y gráficas que se utilicen deberán ser tituladas y numeradas según su orden de aparición. Los títulos han de ser claros, precisos y no repetitivos.

Presentar lo encontrado: patrones o regularidades en un corpus o grupo de producción científica grande (esto puede ser red de coautores, importancia de determinados países en un determinado ámbito, entre otros)

Se obtuvieron 844 citaciones a partir de las dos búsquedas iniciales en Scopus y LILACS. Se excluyeron 33 citaciones obtenidas por la búsqueda en LILACS debido a que eran repetidas o eran tesis de grado. Se tamizaron 811 artículos para evaluar si cumplían con los criterios de selección. Se excluyeron 540 artículos, obteniendo finalmente 271 artículos incluidos en el presente análisis (Figura 2)

PRISMA 2020, propone evaluar Estudios incluidos - Proporcionar el número total de estudios incluidos. - Proporcionar el número total de participantes. - Resumir las características relevantes de los estudios.

Los resultados obtenidos con los indicadores bibliométricos se analizaron y se obtuvieron las conclusiones. Apoyados por herramientas de visualización, por ejemplo VOSviewer.

El presente trabajo mostró el aumento progresivo de publicaciones científicas sobre aplicativos móviles en salud en Latinoamérica desde el 2012. Más de la mitad de los artículos analizados tienen al menos un autor de Brasil. De manera similar, la mayoría de los autores correspondientes identificados eran provenientes de universidades principalmente de Brasil. Finalmente, se evidencia una colaboración en investigación de Brasil con México, y de Brasil con EUA. Hasta nuestro conocimiento, este es el primer análisis bibliométrico de la actividad de investigación publicado en revistas indexadas sobre aplicativos móviles en salud.

Según PRISMA 2020: En cuanto a la discusión propuesta, según PRISMA, se recomienda destacar las “Limitaciones de la evidencia” (Proporcione un breve resumen de las limitaciones de la evidencia incluida en la revisión (por ejemplo, riesgo de sesgo, inconsistencia –heterogeneidad– e imprecisión) y además destacar en la Interpretación (Proporcione una interpretación general de los resultados y sus implicaciones importantes).

Este documento trabaja con datos de Web of Science Core Collection (WoS), Clarivate Analytics. Seleccionamos todos los trabajos indexados que contienen ‘bitcoin’ como tema, lo que hace un total de 1162 documentos, publicados en 703 fuentes (revistas, libros, etc.), durante el período 2012-2019. Estos documentos fueron (co-)escritos por 2293 personas. La gran mayoría de los documentos son de varios autores, siendo sólo 322 documentos de un solo autor. El número medio de autores por documento es de 1,97.

Bitcoin como tema de investigación comprende varias disciplinas. . .

Web of Science asigna artículos indexados a una o más áreas de investigación. Los 1162 artículos considerados en nuestro muestra se asignaron 1543 áreas de investigación. Las cinco principales áreas de investigación se muestran en la Tabla 1 Principales áreas de investigación asignadas a los trabajos de la muestra (Áreas de investigación, Registros, % de 1543).

Tabla 2 Número de artículos publicados por año, con ‘bitcoin’ como tema. Fuente: Web of Science Colección Núcleo (Año, # artículos, Tasa de crecimiento anual)

Indicadores bibliométricos

Con los datos bibliográficos de las publicaciones que resultaron, se aplicarán los siguientes indicadores bibliométricos.

1) Sobre las Áreas de Investigación

Identificar las Áreas de investigación estudiadas. A fin de enfocarse en las áreas que producen la mayor cantidad de artículos. (Tabla 1)

2) Sobre los Años de publicación

Histórico de publicaciones. Este indicador muestra el número de publicaciones realizadas año a año.

Identificar los años de explosión de artículos (Yearly growth rate)

3) Sobre el país de origen de la producción

Distribución geográfica del autor Tabla 3 Los diez principales países de los autores correspondientes (País, Artículos, Frecuencia, Publicaciones de un solo país, Publicaciones de múltiples países, MCP_Ratio)

País de origen. Este presenta el número de publicaciones por país, acorde al país donde se realizó la publicación.

4) Sobre las Citas

Distribución de Citas totales por país Tabla 4 Top ten de citas totales por país (País, Citas totales, Promedio de citas de artículos).

Citación de autores. Este indicador presenta cómo se relacionan y son citados los autores. El método utilizado para encontrar las relaciones y los grupos es el ya nombrado fuerza de asociación, para lo cual se usó también el software VosViewer.

Identificación de la colaboración internacional Además, alrededor del 35% de los artículos publicados son de autores de múltiples países. La colaboración internacional en EE. UU. e India están por debajo del promedio, con una tasa de publicaciones en varios países del 14 % y el 10 %, respectivamente. El internacional la colaboración alcanza un máximo para España, donde el 46% de los trabajos son de este tipo.

5) Sobre la fuente de publicación

Tipo de publicación. Este indicador muestra el número de publicaciones clasificadas por tipo de medio, por ejemplo revista especializada, memoria de conferencia, revisión, libro o capítulo de libro.

Identificar las revistas y libros más importantes sobre el tema. Es importante identificar las fuentes de producción cuando se está haciendo un artículo bibliométrico, porque sería un potencial espacio donde publicar el artículo que se está realizando)

Principales fuentes de publicación Tabla 5 Las diez fuentes más relevantes (Fuentes #Artículos Tipo). Incluir el análisis de las revistas. Tabla 6 (Revista, % Documentos publicados, # Citas, Citas/Documentos, cite Score, Quartil, SIR, % Colaboracion internacional)

6) Sobre los artículos

Artículos muy citados Tabla 7 Artículos altamente citados, en orden descendente por número de citas (Autor (año), Título, Fuente, # Citas).

7) Sobre las Palabras Clave Para las palabras clave se usó el método de fuerza de asociación. Este método busca determinar cuáles categorías se relacionan más entre ellas y se agrupan acorde a ello. Para ello se usó el software libre VosViewer, el cual es un software para análisis bibliométrico que ayuda a crear redes bibliométricas para ver gráficamente estas relaciones.

Palabras clave principales. Tabla 6 Palabras clave principales (Palabras clave del autor (DE), #Artículos, Palabras clave+ (ID), #Artículos)

8) Sobre los autores

Detectar los autores más productivos (con mayor cantidad de artículos) Este muestra el listado de autores y el número de publicaciones científicas de cada uno. Tabla 8 Autores más productivos (Autores, Institución, # artículos).

Índice de concentración de autores, como una especie de entropía, definido en el artículo publicado en IDEAS.

Grado de concentración de las variables seleccionadas Propuesta en el artículo ...

En este subapartado analizamos globalmente varias variables bibliométricas, con el fin de mostrar el grado de concentración de las mismas. Una característica importante en los estudios bibliométricos es la uniformidad de la contribución de los autores, países y revistas a una investigación. La teoría de la información proporciona métricas alternativas a las medidas estadísticas tradicionales de concentración, como la desviación estándar, asimetría o curtosis. En particular, Shannon y Weaver (1949) desarrollaron la célebre entropía de Shannon. Dada una distribución de probabilidad de tributación

Tabla 9 Índice de concentración entrópica () de seleccionados variables (Variable, H)

Tabla 10 Distribución observada del número de autores que escribieron un número determinado de artículos y valores ajustados por la ley de Lotka (# Artículos, # Autores, Frecuencia observada, Frecuencia ajustada)

9) Sobre las organizaciones a las que pertenecen los autores

Afilación de autores. Este muestra las organizaciones a las cuales pertenecen los autores y el número de publicaciones asociadas a cada organización.

10) Sobre los patrocinadores

Patrocinador Este indica los patrocinadores que financiaron las investigaciones que generaron las publicaciones. El número corresponde a las publicaciones que tiene asociadas cada patrocinador.

Análisis de las visualizaciones

- 1) Enfocados en los artículos. Representando los artículos por nodos, junto con el tamaño del nodo representando la cantidad de citas que recibe ese nodo. El nodo se etiqueta con la palabra clave más característica. Análisis de las visualizaciones (por ejemplo, en el artículo de bitcoin se observó que hay tres cluster importantes, uno más vinculado a business economics, otra que era mas computer science pero divididos en una parte que se dedicaba a blockchain duro (más teórico) y una parte más aplicada a los protocolos que se aplicaban a criptomonedas en concreto y bitcoin en particular).
- 2) Enfocados en los autores. Donde se destaque los autores y coautores mediante citas.

Gráficos de citas, fuentes y autores Las siguientes cifras fueron generadas utilizando el software VOS viewer, que permite contar las palabras que aparecen en el título, resumen y palabras clave para construir todas las relaciones que aparecen entre los diferentes documentos publicados en la Web of Science (van Eck y Waltman, 2010).

Fig. 1. Mapa de nubes de palabras en títulos y resúmenes (conteo completo), generado con VOSviewer (<http://www.vosviewer.com/>). (Para interpretación de las referencias al color en el texto, se remite al lector a la versión web de este artículo).

Fig. 2. Mapa de nubes de palabras en títulos y resúmenes (conteo binario), generado con VOSviewer (<http://www.vosviewer.com/>). (Para interpretación de las referencias al color en el texto, se remite al lector a la versión web de este artículo).

Fig. 3. Mapa en la nube de revistas donde se publicaron artículos sobre bitcoin, generado con VOSviewer (<http://www.vosviewer.com/>).

Fig. 4. Mapa en la nube de revistas de autores con artículos sobre bitcoin publicados, generado con VOSviewer (<http://www.vosviewer.com/>).

Consideraciones éticas

El presente estudio es de tipo bibliométrico, que incluye un análisis de artículos publicados en revistas científicas indexadas en la base de datos Scopus y LILACS, por lo cual no fue necesario solicitar el consentimiento a los autores para analizar y presentar la información. Adicionalmente, no se obtuvo, analizó, ni manipuló información sensible de personas ni animales.

Potenciales líneas de investigación

La figura 2 señala una clara separación entre los artículos orientados a la tecnología y los orientados a la economía. Teniendo en cuenta la tecnología intrínseca componente nológico de bitcoin y su fuerte impacto económico, detectamos una falta de trabajos interdisciplinarios.

...

Finalmente, se necesita investigación sobre los aspectos de comportamiento de bitcoin, para comprender un mercado, donde hay heterogéneos individuos, interactuando en tiempo real. En relación con este aspecto, casi no hay documentos sobre el impacto social de bitcoin.

Conclusiones

Resultados destacados y Próximos pasos.

- Se publicaron 271 artículos sobre aplicativos móviles en salud con filiación de algún país latinoamericano desde el 2012 hasta el 2019.
- Latinoamérica se encuentra en fase de crecimiento exponencial de artículos sobre aplicativos móviles en salud asimismo se encontró un incremento progresivo de publicaciones en revistas Q1.
- El mayor grado de colaboración en investigación se encontró entre Brasil y México, y entre Brasil y EUA.

Conclusiones

Este estudio muestra que la literatura de las criptomonedas comprende principalmente una combinación de informática y economía. a pesar de que es originalmente un producto tecnológico, las principales aplicaciones de blockchain son las criptomonedas. Entre ellos, bitcoin es el actor dominante, tanto en el mercado como en el interés literario. El número de documentos publicados sobre este tema ha sido aumentando a una tasa anual del 124%, aunque disminuyendo a lo largo de los años. Este artículo constituye la primera revisión bibliométrica completa estudio de la literatura bitcoin, que comprende todos los artículos indexados en la Web of Science desde 2012. La gran cantidad de datos (1162 papers) permite encontrar resultados significativos con respecto a los principales académicos, principales revistas y palabras clave de este campo de investigación multidisciplinario. Detectamos una alta concentración en los países editores. Sin embargo, los autores son diversos y menos concentrados que en los principales diarios de finanzas. Además, las citas se concentran en unos pocos artículos. En futuros trabajos nos gustaría estudiar el temporal evolución de las palabras clave. Además, nos gustaría probar si el rápido crecimiento y la posterior caída del precio de bitcoin en 2017 han afectó la investigación relacionada con bitcoin. Finalmente, las palabras relacionadas como “criptomonedas”, “fintech” y “préstamos entre pares” serán abordado en investigaciones posteriores.

Expresiones de gratitud

Los autores agradecen al editor y a un revisor anónimo por sus útiles comentarios. Ignasi Merediz-Solà reconoce la apoyo a una beca de investigación de la Secretaría de Estado de Educación y Formación Profesional (España) # 2018 COLAB 00083 Fig. 4. Mapa en la nube de revistas de autores con artículos sobre bitcoin publicados, generado con VOSviewer (<http://www.vosviewer.com/>). I. Merediz-Solà y A.F. Bariviera Research in International Business and Finance 50 (2019) 294–305 303 (BOE núm. 194, 11.08.2018).

Referencias o Referencias bibliográficas

Bibliografía

Clarivate Analytics, 2018a. InCites Indicators Handbook. Clarivate Analytics, Philadelphia (PA) (accessed 20.02.19). Clarivate Analytics, 2018b. Web of Science. Trust the Difference. (accessed 18.02.19). <http://wokinfo.com/>. Cuccurullo, C., Aria, M., Sarto, F., 2016. Foundations and trends in performance management. A twenty-five years bibliometric analysis in business and public administration domains. *Scientometrics* 108, 595–611. Aria, M., Cuccurullo, C., 2017. bibliometrix: an r-tool for comprehensive science mapping analysis. *J. Inform.* 11, 959–975. Broadus, R.N., 1987. Toward a definition of “bibliometrics”. *Scientometrics* 12, 373–379. Pritchard, A., 1969. Statistical bibliography or bibliometrics? *J. Doc.* 25, 348–349.

Apéndices

...