



Tendencias del big data y cloud computing:

Bibliometría del 2010 al 2020

Jesús Hernández Sánchez

jesus.hernandez07@upaep.edu.mx

Erick Leobardo Álvarez Aros

erickleobardo.alvarez@upaep.mx

Rosa María Cantón Croda

rosamaria.canton@upaep.mx

Universidad Popular Autónoma del Estado de Puebla (UPAEP)
Decanato de Ingeniería y Negocios, Posgrados de Planeación Estratégica y
Dirección de Tecnología, calle 17 Sur 711, Barrio de Santiago, C.P. 72410
Puebla, Puebla,

RESUMEN

En el presente estudio se identificaron las tendencias más significativas de los documentos científicos de alto impacto analizados con respecto al Big Data y Cloud Computing durante el periodo comprendido entre los años 2010 al 2020, cuya revisión se realizó en las bases de datos Web of Science (WoS) y Scopus de 111 artículos. Los resultados fueron varios, como, por ejemplo, B. Dong como el autor con más publicaciones, China, Estados Unidos e India como los países con más estudios y estos primeros los que más colaboran entre si; por mencionar algunos.

Palabras clave: Big Data, Cloud Computing, Bibliometrix, ciencia de datos, tecnología, software R.

Big data and cloud computing trends: Bibliometrics from 2010 to 2020

ABSTRACT

This study identified the most significant trends in the high impact scientific documents analyzed with respect to Big Data and Cloud Computing during the period between 2010 and 2020, whose review was carried out in the Web of Science databases (WoS) and Scopus of 111 articles. The results were various, such as, for example, B. Dong as the author with the most publications, China, the United States, and India as the countries with the most studies and the first the most collaborative among themselves; to name a few. The following.

Keywords: Big Data, Cloud Computing, Bibliometrix, data science, technology, R software.

Artículo recibido: 10. Mayo. 2021

Aceptado para publicación: 28. junio. 2021

Correspondencia: jesus.hernandez07@upaep.edu.mx

Conflictos de Interés: Ninguna que declarar

I. INTRODUCCIÓN

El presente trabajo considera uno de los conceptos más conocidos, estudiados y analizados mediante múltiples nombres referidos en investigaciones anteriores, los términos comúnmente encontrados son “datos masivos”, “macro datos”, “inteligencia de datos” o “datos a gran escala”, pero para fines de esta investigación se usará Big Data para referirnos a la misma tecnología el resto del estudio.

El objetivo de este estudio fue identificar las tendencias al realizar una revisión bibliométrica, es decir, se obtuvieron datos de los artículos y autores más sobresalientes del tema de Big Data y Cloud Computing proporcionando definiciones, características y clasificación de estas tecnologías, al mismo tiempo, se detallan algunas discusiones sobre la computación en la nube. También se aborda la relación entre ambas tecnologías, así como los sistemas de almacenamiento de datos. La justificación al realizar este trabajo es debido a la gran importancia que está teniendo en la actualidad estas tecnologías, en donde cualquier persona con o sin conocimientos tecnológicos tienen cierto grado de interés de las formas en que las empresas almacenan la información que se genera desde Facebook, Twitter o Instagram, por ello el Big Data alcanza muchos ámbitos de nuestra vida desde la economía, educación, ciencias tecnológicas, deportes, seguridad, política, entre otros.

En cuanto a la metodología utilizada para la revisión bibliométrica, se inició con la búsqueda de artículos relacionados directamente con el tema en las bases de datos WoS, Scopus y Google Scholar, apoyándose en estudios similares, posteriormente se revisaron 111 artículos con alto grado de impacto con el propósito de obtener el mayor número de datos, cifras, archivos de apoyo, gráficas, tablas, así como diversas opiniones de expertos en la materia.

En cuanto al contenido del marco teórico, la información se obtuvo de materiales académicos, formando un puzzle lo más homogéneo que permita obtener una visión del concepto y las influencias en el mismo. Posteriormente se llevó a cabo la revisión de la literatura que consistió en profundizar el tema de Big Data y Cloud Computing de tal forma de obtener información confiable, veraz, así como de suma relevancia.

Después de lo anterior se obtiene los siguientes apartados que constituyen esta investigación: En primer lugar, un acercamiento a la conceptualización de Big Data y

Cloud Computing para establecer las bases de la investigación, en segundo lugar, el almacenamiento local y remoto para comprender la importancia de los sistemas de Cloud Computing, en tercer lugar, los problemas de seguridad de la información y posibles soluciones, en cuarto lugar, los costos de almacenamiento de datos así como la disponibilidad, por último algunas plataformas consideradas casos de éxito en esta temática.

II. MARCO TEÓRICO

Según Hashem, Yaqoob, Anuar, Mokhtar, Gani y Khan (2015) establecen que abordar el tema de Big Data es una tarea desafiante que requiere mucho tiempo de estudio, así como una gran infraestructura computacional para garantizar el procesamiento y análisis de datos exitosos. Por otro lado, la computación en la nube es considerada una tecnología poderosa para realizar computación compleja, a gran escala que elimina la necesidad de mantener hardware informático costoso, espacio dedicado y software.

Actualmente la velocidad de generación, así como el crecimiento de datos está aumentando debido a la proliferación de dispositivos móviles y otros sensores conectados a Internet, estos datos brindan oportunidades que permiten a las empresas de todas las industrias obtener información empresarial en tiempo real. Según Yang, Huang, Li, Liu y Hu (2017) establecen que el Big Data ha surgido en los últimos años como un nuevo paradigma que proporciona abundantes datos, así como oportunidades para mejorar y permitir la investigación en aplicaciones de soporte de decisiones con un valor sin precedentes para el mundo digital, aplicaciones que incluyen negocios, ciencias e ingeniería.

De acuerdo con Cai, Xu, Jiang y Vasilakos (2016) las aplicaciones relacionadas con Internet de las cosas (IoT) se han convertido en un campo importante tanto para ingenieros como para investigadores, lo que refleja la magnitud y el impacto de los problemas relacionados con los datos que deben resolverse en las organizaciones empresariales contemporáneas, especialmente en la computación en la nube (Stergiou & Psannis, 2017).

Según Sookhak (2015) la computación en la nube se ha convertido en un nuevo paradigma que ofrece un gran potencial para almacenar datos de forma remota. Por ello, muchas organizaciones han reducido la carga del almacenamiento y

mantenimiento de datos locales al externalizar el almacenamiento de datos a la nube. Sin embargo, la integridad y seguridad de los datos subcontratados sigue siendo motivo de gran preocupación para los propietarios de datos debido a la falta de control, así como posesión física sobre los datos.

III. REVISIÓN DE LA LITERATURA

El aumento continuo del volumen, así como detalles de los datos capturados por las organizaciones, como el auge de las redes sociales, Internet de las Cosas (IoT) y multimedia, han producido un flujo abrumador de datos en formato estructurada y no estructurada. La creación de datos está ocurriendo a una velocidad récord, referida aquí como Big Data, y se ha convertido en una tendencia ampliamente reconocida (Zhang, Yang, Chen, Li & Bu, 2018).

Según Hashem et al. (2018) estudiar el tema de Big Data es una tarea retadora, que requiere mucho tiempo y una gran infraestructura computacional para garantizar un procesamiento y análisis de datos exitosos, por tal situación dicho tema está llamando poderosamente la atención de la academia, el gobierno y la industria. Así mismo Yang et al. (2017) aborda que esta tecnología ha surgido en los últimos años como un nuevo paradigma que proporciona abundantes datos y oportunidades para mejorar y permitir la investigación y aplicaciones de soporte de decisiones con un valor sin precedentes para el mundo digital, aplicaciones que incluyen negocios, ciencias e ingeniería.

Conceptualización de Big Data y Cloud Computing

Según Hashem et al., (2018) establece que el término Big Data se caracteriza por cuatro V's: volumen, variedad, velocidad y veracidad, por lo tanto, la siguiente definición se propone con base a las características mencionadas anteriormente, así como el análisis de la esencia de los grandes datos. Big Data es un conjunto de técnicas y tecnologías que requieren nuevas formas de integración para descubrir grandes valores ocultos de grandes conjuntos de datos en una escala diversa, compleja y de gran tamaño (Manogaran, Thota, & Kumar, 2016).

En el mismo sentido se aborda el término computación en la nube denominada Cloud Computing para usos posteriores, la cual es una tecnología poderosa para realizar computación compleja y a gran escala que elimina la necesidad de mantener hardware informático costoso, espacio dedicado y software (Hashem et al., 2015).

Para Sookhak (2015) el Cloud Computing se ha convertido en un nuevo paradigma que ofrece un gran potencial para almacenar datos de forma remota, es por ello que actualmente, muchas organizaciones han reducido la carga del almacenamiento y mantenimiento de datos locales al externalizar el almacenamiento de datos a la nube. Sin embargo, la integridad y seguridad de los datos subcontratados sigue siendo motivo de gran preocupación para los propietarios de datos debido a la falta de control y posesión física sobre los datos, por lo cual algunos investigadores han propuesto técnicas de auditoría remota de datos (RDA).

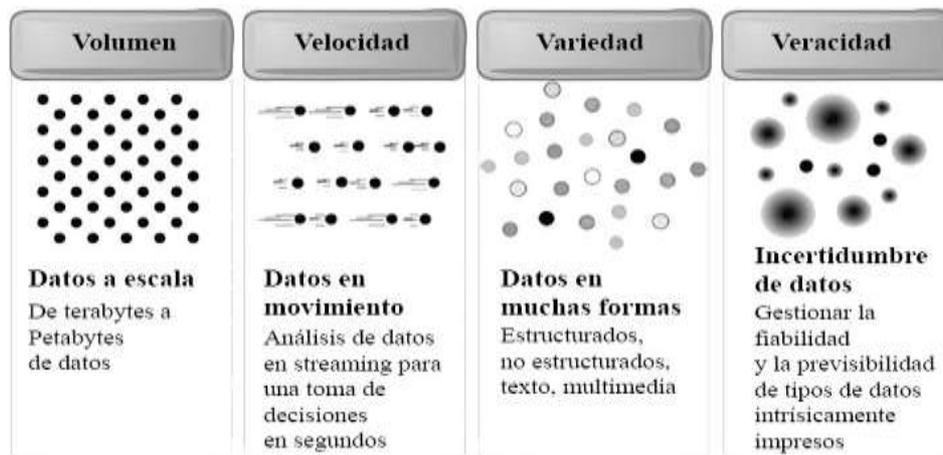


Figura 1. Dimensiones o características de Big Data

En este sentido se propone un marco seguro basado en el Cloud Computing para la gestión de la información de Big Data en redes inteligentes, que se denomina "marco inteligente", la idea principal de este marco es construir una estructura jerárquica de centros de computación en la nube para proporcionar diferentes tipos de servicios informáticos para gestión de información y análisis de Big Data, como resultado, este marco propuesto logra no solo escalabilidad y flexibilidad, sino también características de seguridad en donde se han implementado pruebas de concepto para este marco con una gestión simple basada en la identidad para la confidencialidad de los datos (Baek, Vu, Liu, Huang & Xiang, 2014)

Almacenamiento local y remoto

La generación de datos a gran escala y velocidad ha ocasionado la búsqueda de nuevos mecanismos de almacenamiento de información, por ello, la computación en la nube se ha convertido en una importante alternativa que ofrece un gran potencial para almacenar datos de forma remota. En consecuencia, muchas organizaciones han decidido reducir la carga del almacenamiento y mantenimiento de datos locales al

externalizar el almacenamiento de datos en la nube (Sookhak, 2015).

Por otro lado, el almacenamiento local tiene sus beneficios en la imposibilidad de que los datos salgan de nuestro disco si disponemos de un sistema con la seguridad apropiada, tiene un costo medio y particularmente se debe procurar que el dispositivo tenga un sistema de backup (respaldo) apropiado, la gran desventaja de este sistema es que no se tiene los datos disponibles en cualquier lugar, si no, solamente en la red de trabajo (Wang, Ma, Yan, Chang & Zomaya, 2018).

Por ello, es importante mencionar que las tecnologías han evolucionado para atender las necesidades de almacenamiento, actualmente se encuentran en una etapa compleja donde es más fácil producir datos que guardarlos y administrarlos. Dicho lo anterior, se tiene que cada año en el mundo se produce entre tres y cinco Exabytes de información (Lyman & Varian, 2003). Aunque en 2007 se almacenaron de manera comprimida 2.9 Zettabytes, se comunicaron cerca de dos Zettabytes y se ejecutaron 6.4 Exabytes de instrucciones por segundo en computadoras de propósito general (Hilbert & López, 2011), se cree que en el año 2020 se tendrán 35 Zettabytes (Li & Cao, 2014).

IV. METODOLOGÍA

En este apartado de metodología se realizó un análisis bibliométrico considerado también como un estudio bibliométrico de la cantidad y calidad de artículos basados en el tema de Big Data y Cloud Computing, así como la relación existente entre los mismos, para ello se ha llevado a cabo la selección y búsqueda de dichos artículos en las principales bases de datos, así como buscadores de Internet.

A partir de las búsquedas realizadas el 30 y 31 de enero de 2020 acerca de los artículos más importante o sobresalientes en las bases de datos WoS y Scopus y en el buscador google académico, se obtienen como resultado archivos en diferentes formatos como son: archivos de texto (txt), archivos en formato comma-separated values (CSV) y archivos en formato BibTex (Bib).

Según Aria y Cuccurullo (2017) en el entorno R, se han publicado recientemente otros paquetes en el repositorio oficial (CRAN, The Comprehensive RArchive Network, <https://cran.r-project.org/>) que abordan la bibliometría. Cada uno de ellos proporciona funciones de análisis específicas; sin embargo, ninguno aborda todo el flujo de trabajo.

Es importante indicar que durante el proceso de búsquedas se utilizaron diferentes operadores booleanos, esto con la finalidad de obtener mejores resultados, entre las búsquedas más sobresalientes fueron: Big Data, Big Data AND Cloud Computing, Cloud Computing, Industry 4.0 AND Big Data, Industry 4.0 AND Cloud Computing. El resultado obtenido de WoS se puede exportar y analizar fácilmente para la mayoría de los autores, instituciones, países, revistas, áreas temáticas, citas y mapeo de palabras clave más activos. Las palabras clave utilizadas en este estudio son relevantes para llegar a resultados éxitos (Belfiore, Iovino y Tafuri, 2019). Después de las búsquedas, se realiza una criba en la que, mediante la lectura de los resúmenes de cada artículo, se eligen los artículos más importantes y de interés particular, dando como resultado una lista más concreta de los artículos a revisar con más detalle, teniendo como resultado un total de 111 artículos para su análisis, exploración y determinación de los datos más sobresalientes respecto al tema investigado.

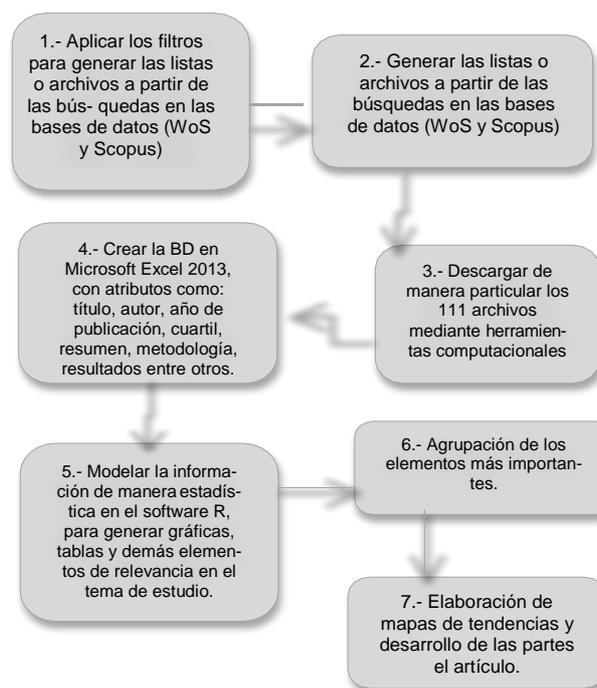


Figura 2. Diagrama de la metodología

Un análisis bibliométrico usa el paquete bibliometrix en el software R, este paquete utiliza los metadatos en las citas de la Web of Science para calcular y clasificar la producción del país, las fuentes de las revistas y las colaboraciones de los países (Nafade, Nash, Huddart, Pande, Gebreselas-sie, Lienhardt & Pai, 2018). Para hacer un manejo óptimo,

eficiente y adecuado del contenido de los artículos, se realiza una búsqueda web particular de cada uno de los 111 artículos, para lo cual utilizamos herramientas computacionales para acceder al documento digital original o copia, y en consecuencia descargar dichos artículos (en su idioma original o traducido), así como tener información completa para su posterior estudio y desarrollo. En referencia al contenido del marco teórico, toda la información se alimenta de los materiales académicos descargados, teniendo como resultado una visión más completa del Big Data y el Cloud Computing, así como los retos y oportunidades del mismo en el mundo digital.

De tal forma que se lleva a cabo la metodología de recogida de datos y utilización de los mismos, con la finalidad de poseer los elementos de información con mayor credibilidad respecto al tema investigado. Después del proceso anterior se tiene como resultado los siguientes apartados incluidos en el trabajo de investigación. En primer lugar, se desarrolla un breve resumen, la revisión del marco teórico, la revisión de la literatura, el planteamiento de la metodología, análisis de datos y resultados, y por último las conclusiones. Enseguida se presenta el diagrama de la metodología aplicada en el desarrollo de esta investigación. Es importante mencionar que para la obtención de datos confiables se revisaron cada uno de los cuartiles a los que pertenecen las revistas responsables de publicar los 111 artículos del tema en cuestión, en donde se indica que la mayoría son del Q1, Q2, Q3 y muy pocos de Q4.

V. RESULTADOS

La computación en la nube es una tecnología poderosa para realizar computación compleja y a gran escala que elimina la necesidad de mantener hardware informático costoso, espacio dedicado y software sofisticado. Además, el Big Data es un término utilizado para referirse al aumento en el volumen de datos que son difíciles de almacenar, procesar y analizar a través de las tecnologías de bases de datos tradicionales (Hashem et al., 2015).

El tema de Big Data y Cloud Computing ha tenido un gran auge en los últimos años, de acuerdo con la revisión bibliométrica en Scopus y WoS de 111 artículos, en el Figura 3 se muestran los autores más relevantes que estudian estos temas, dichos autores son: Dong F, Hsu Ch, Agrawal D y Li K.

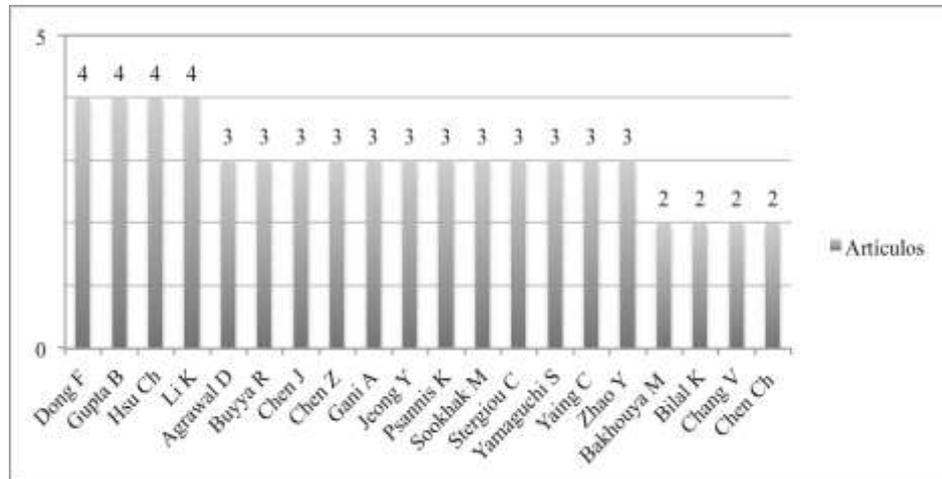


Figura 3. Principales autores de artículos sobre el Big Data y Cloud Computing

De acuerdo a la Figura 3 se observa que 276 autores han escrito al menos un artículo, 28 autores han desarrollado estudios de al menos dos artículos, 12 autores han escrito tres artículos y solo 4 autores han escrito cuatro artículos. Por otro lado, Gani A. (2015) fue uno de los autores más citados con un total de 741 citas, seguido de Anuar Nb. (2015) con un total de 681 citas y en tercer lugar encontramos a Has-hemlat (2015) con el mismo número de 681 citas.

Además, se examinó la producción científica de los autores a lo largo del tiempo, en la Figura 4 se muestra que Gupta B. en el año 2018 desarrollo tres artículos relacionados con este tema.

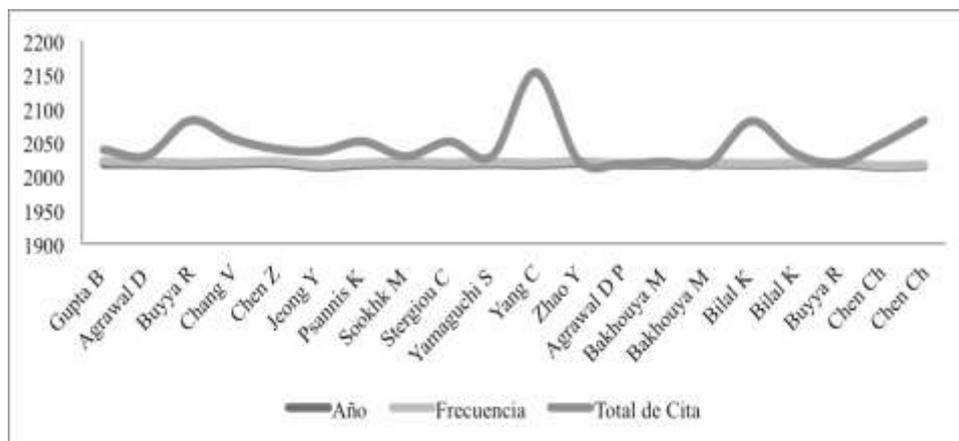


Figura 4. Principales autores de artículos sobre el Big Data y Cloud Computing

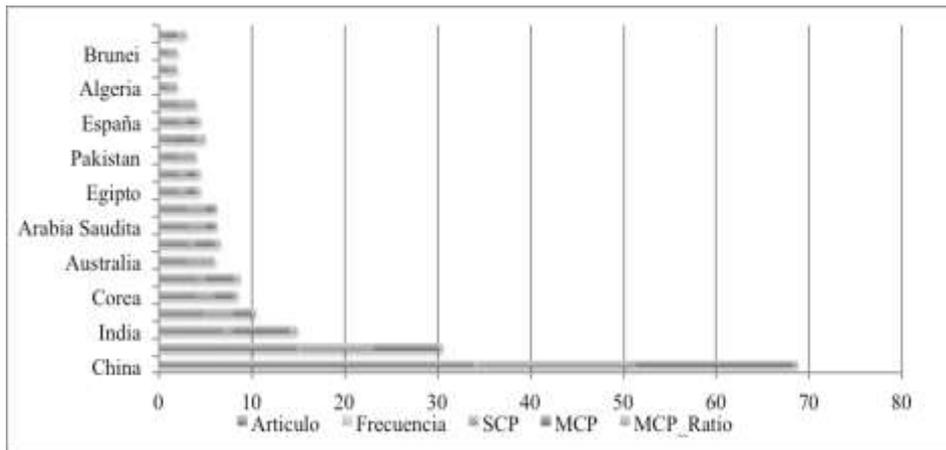


Figura 5. Principales países y la producción científica de artículos sobre el *Big Data* y *Cloud Computing*

La producción científica es uno de los parámetros más importantes para predecir el desarrollo de un país, por ello en la Figura 5 se observa que China es uno de los países que encabeza los estudios del tema de Big Data y Cloud Computing, seguidos de Estados Unidos, India y Grecia.

Sin embargo, de acuerdo a la Figura 6, se observa que Malasia encabeza el ranking de los países más citados en dicho tema con 741 citas totales y 185.25 citas promedios por artículo, seguido de Estados Unidos y China, cabe mencionar que los países de América Latina no tienen grandes estudios relacionados al tema de Big Data y Cloud Computing, por ello podemos encontrar un vacío en la literatura para abordar dichos temas en diferentes áreas de conocimiento.

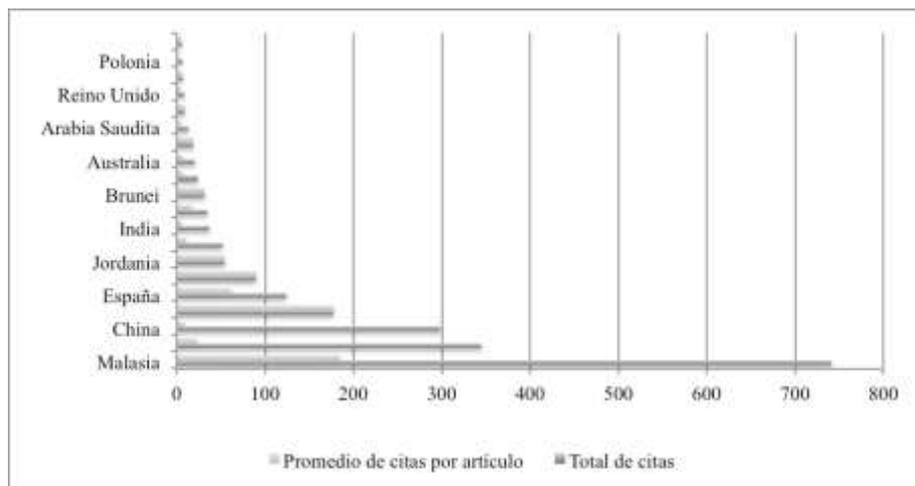


Figura 6. Principales países más citados de artículos sobre el *Big Data* y *Cloud Computing*

En la Figura 8 se muestra los autores más citados, en donde se aprecia que Cuzzocrea, Yang y Sultan son los tres autores más citados, en este sentido es importante señalar que mediante este ejercicio se puede apreciar a los investigadores que se debe considerar para plantear algunas situaciones de estudio. Por otro lado, al realizar búsquedas de información relacionada con Big Data y Cloud Computing encontrará las co-citaciones que unos autores hacen de otros, por lo tanto, existe una relación directa de la temática en el desarrollo de investigaciones. Esta revisión presenta innovaciones futuras y una agenda de investigación para la computación en la nube que respalda la transformación del volumen, la velocidad, la variedad y la veracidad en valores de Big Data para aplicaciones y aplicaciones de ciencias de la tierra digitales locales a globales (Yang et al., 2017).

En la Figura 9 Revistas con más publicaciones de Big Data, se observa el ranking de las revistas que más publican artículos relacionados con estas tecnologías, es importante mencionar que con el surgimiento continuo de una variedad de nuevos métodos de difusión de información y el surgimiento de las tecnologías de computación en la nube e IoT, los datos aumentan constantemente con una alta velocidad, por ello la escala de datos globales aumenta continuamente a un ritmo de dos veces cada dos años (Chen, Li, Tang, Bilal, Yu, Weng & Li, 2016).

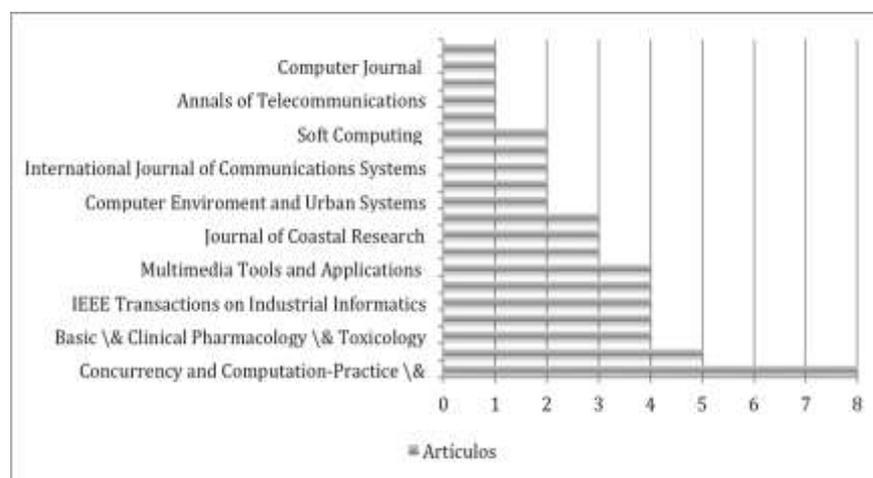


Figura 9. Revistas con más publicaciones de Big Data y Cloud Computing

Según Baek et al. (2014) La computación en la nube se ha vuelto popular recientemente debido a varias ventajas sobre los modelos informáticos tradicionales. Las ventajas típicas incluyen flexibilidad, escalabilidad, agilidad, eficiencia energética y ahorro de costos. Por lo anterior, al momento de implementar nuevos

sistemas informáticos aparecen una serie de palabras que convergen en un mapa semántico y por ende permiten la creación de nuevos espacios al conocimiento.

VI. CONCLUSIONES

Conforme al objetivo de este trabajo, se identificaron las tendencias más significativas de los documentos científicos de alto impacto analizados con respecto al Big Data y Cloud Computing durante el periodo comprendido entre los años 2010 al 2020.

Los datos brindan oportunidades que permiten a las empresas de todas las industrias obtener información comercial en tiempo real, es así como con el uso de servicios en la nube para almacenar, procesar y analizar datos que ha estado disponible durante algún tiempo obtienes enormes beneficios, por un lado, ha cambiado el contexto de la tecnología de la información y ha hecho realidad las promesas del modelo de servicio bajo demanda (Hashem, 2015).

La presente investigación se enfocó en comprender la relación existente entre el Big Data y Cloud Computing, así como la aplicación de dichas tecnologías en las actividades cotidianas de las empresas, es importante mencionar que dentro de las limitaciones de esta investigación se encuentra la búsqueda de información en bases de datos WoS, Scopus y algunos artículos de google Scholar, por lo cual se recomienda incluir para futuras investigaciones algunas bases de datos universitarias, así como usar otros criterios para ampliar la búsqueda considerando las palabras claves “Big Data*” y “Cloud Computing and applications”.

Los resultados del análisis bibliométrico demuestran la importancia que tiene hoy en día la investigación sobre el Big Data y Cloud Computing, siendo notable el incremento a partir del año 2010, por otro lado, es importante señalar que los países más destacados en la investigación de estos temas son: China, Estados Unidos, Japón, Malasia, Australia seguidos por otros países asiáticos. Cabe resaltar que, en la región de Latinoamérica, los temas de Big Data y Cloud Computing son poco estudiados.

En este estudio bibliométrico se reconoce el trabajo de los principales autores y países que realizan los mejores estudios de estos temas, teniendo en cuenta que Big Data es una nueva tecnología que se desarrolla rápidamente en diversos campos de estudio, hemos intentado combinar con el Cloud Computing, esto para comprender su importancia en la actualidad y verificar las características comunes, así como

descubrir los beneficios de estas dos tecnologías con respecto a su uso con las aplicaciones de Big Data (Lo' ai, 2016).

Una de las limitantes respecto a estas tecnologías es el rezago tecnológico de los países en desarrollo, así como el acceso limitado a las bases de datos como WoS y Scopus. Es importante recalcar que hay muchos retos y problemas por resolver, así como oportunidades para cambiar el uso de las tecnologías de Big Data y Cloud Computing en el futuro.

VII. AGRADECIMIENTOS

Expreso mi agradecimiento al Consejo Nacional de Ciencia y Tecnología (CONACyT) por el apoyo en el financiamiento otorgado para la realización de estudios doctorado en el programa de Planeación Estratégica y Dirección de Tecnología en la UPAEP en el período del 2020 al 2023.

VIII. REFERENCIAS

- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of In- formetrics*, 11(4), 959-975.
- Baek, J., Vu, Q. H., Liu, J. K., Huang, X., & Xiang, Y. (2014). A secure cloud computing based framework for big data information management of smart grid. *IEEE transac- tions on cloud computing*, 3(2), 233-244. doi: 10.1109/TCC.2014.2359460
- Belfiore, P., Iovino, S., & Tafuri, D. (2019). Sport manage- ment and educational management: a bibliometric analysis. *Sport Science*, 12(1), 61-64.
- Cai, H., Xu, B., Jiang, L., & Vasilakos, A. V. (2016). IoT- based big data storage systems in cloud computing: perspectives and challenges. *IEEE Internet of Things Journal*, 4(1), 75-87. DOI: 10.1109/JIOT.2016.2619369
- Chen, J., Li, K., Tang, Z., Bilal, K., Yu, S., Weng, C., & Li, K. (2016). A parallel random forest algorithm for big data in a spark cloud computing environment. *IEEE Transac- tions on Parallel and Distributed Systems*, 28(4), 919- 933. DOI: 10.1109/TPDS.2016.2603511
- Gai, K., Qiu, M., Zhao, H., & Xiong, J. (2016, June). Privacy- aware adaptive data encryption strategy of big data in cloud computing. In *2016 IEEE 3rd International Confe- rence on Cyber Security and Cloud Computing (CSCloud)* (pp. 273-278). IEEE. doi: 10.1109/CSCloud.2016.52

- Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information systems*, 47, 98-115. DOI: 10.1016/j.is.2014.07.006
- Hilbert, M. y López, P. (2011). The world's technological capacity to store, communicate, and compute information. *Science*, 332(60), 60-65.
- Lo'ai, A. T., Mehmood, R., Benkhelifa, E., & Song, H. (2016). Mobile cloud computing model and big data analysis for healthcare applications. *IEEE Access*, 4, 6171-6180. DOI: 10.1109/ACCESS.2016.2613278
- Lo'ai, A. T., Bakheder, W., & Song, H. (2016, June). A mobile cloud computing model using the cloudlet scheme for big data applications. In *2016 IEEE First International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)* (pp. 73-77). IEEE. DOI: 10.1109/CHASE.2016.40
- Li, M. y Cao, S. (2014). A serie method of massive information storage, retrieval and sharing. En *Mechatronics and Automation (ICMA), 2014 IEEE International Conference on* (pp. 1171-1175). IEEE.
- Li, J., Huang, L., Zhou, Y., He, S., & Ming, Z. (2017). Computation partitioning for mobile cloud computing in a big data environment. *IEEE Transactions on Industrial Informatics*, 13(4), 2009-2018. DOI: 10.1109/TII.2017.2651880
- Li, Y., Gai, K., Qiu, L., Qiu, M., & Zhao, H. (2017). Intelligent cryptography approach for secure distributed big data storage in cloud computing. *Information Sciences*, 387, 103-115. DOI: 10.1016/j.ins.2016.09.005
- Liu, X., Singh, P. V., & Srinivasan, K. (2016). A structured analysis of unstructured big data by leveraging cloud computing. *Marketing Science*, 35(3), 363-388. DOI: 10.1287/mksc.2015.0972
- Lyman, P. y Varian, H. (2003). *How much information*. Estados Unidos: Universidad de California. Recuperado de <http://www2.sims.berkeley.edu/research/projects/howmuchinfo2003>
- Manogaran, G., Thota, C., & Kumar, M. V. (2016). Meta- CloudDataStorage architecture for big data security in cloud computing. *Procedia Computer Science*, 87, 128-133. DOI: 10.1016/j.procs.2016.05.138

- Nafade, V., Nash, M., Huddart, S., Pande, T., Gebreselasie, N., Lienhardt, C., & Pai, M. (2018). A bibliometric analysis of tuberculosis research, 2007–2016. *PLoS One*, 13(6).
- Pérez, M. M. S. (2015). Big Data O La Acumulación Masiva De Datos Sanitarios : Derechos En Riesgo, 10–12.
- Sookhak, M. (2015). Dynamic remote data auditing for securing big data storage in cloud computing (Doctoral dissertation, University of Malaya). DOI: 10.1016/j.ins.2015.09.004
- Stergiou, C., & Psannis, K. E. (2017). Recent advances delivered by Mobile Cloud Computing and Internet of Things for Big Data applications: a survey. *International Journal of Network Management*, 27(3), e1930. DOI: 10.1002/nem.1930
- Varatharajan, R., Manogaran, G., & Priyan, M. K. (2018). A big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing. *Multimedia Tools and Applications*, 10195-10215. DOI:10.1007/s11042-017-5318-1
- Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2017). Big Data and cloud computing: innovation opportunities and challenges. *International Journal of Digital Earth*, 10(1), 13- 53.
- Wang, L., Ma, Y., Yan, J., Chang, V., & Zomaya, A. Y. (2018). pipsCloud: High performance cloud computing for remote sensing big data management and processing. *Future Generation Computer Systems*, 78, 353- 368. DOI: 10.1016/j.future.2016.06.009
- Xu, J., Huang, E., Chen, C. H., & Lee, L. H. (2015). Simulation optimization: A review and exploration in the new era of cloud computing and big data. *Asia-Pacific Journal of Operational Research*, 32(03), 1550019. DOI:10.1142/S0217595915500190
- Yang, C., Huang, Q., Li, Z., Liu, K., & Hu, F. (2017). Big Data and cloud computing: innovation opportunities and challenges. *International Journal of Digital Earth*, 10(1), 13-53.
- Zhang, Q., Yang, L. T., Chen, Z., Li, P., & Bu, F. (2018). An adaptive dropout deep computation model for industrial IoT big data learning with crowdsourcing to cloud computing. *IEEE Transactions on Industrial Informatics*, 15(4), 2330-

2337. DOI: 10.1109/TII.2018.2791424