

Estudio de la Realidad y Viabilidad de la Formación en BigData en la Academia Chilena

Cristian L. Vidal-Silva^{(1)*}, Erika A. Madariaga⁽²⁾, José M. Rubio⁽³⁾ y Luis A. Urzúa⁽⁴⁾

(1) Ingeniería Civil Informática, Escuela de Ingeniería, Universidad Viña del Mar, Campus Rodelillo, Agua Santa 7055, Viña del Mar - Chile. (e-mail: cristian.vidal@uvm.cl)

(2) Escuela de Computación e Informática, Facultad de Ingeniería, Ciencia y Tecnología, Universidad Bernardo O'Higgins, Avenida Viel 1497, Ruta 5 Sur, Santiago - Chile. (e-mail: erika.madariaga@ubo.cl)

(3) Área Académica de Informática y Telecomunicaciones, Universidad Tecnológica de Chile INACAP, Av. Vitacura 10.151, Vitacura, Santiago - Chile. (e-mail: jrubiol@inacap.cl)

(4) Escuela de Kinesiología, Facultad de Salud, Universidad Santo Tomás, Talca - Chile. (e-mail: lurzua@santotomas.cl)

* Autor a quien debe ser dirigida la correspondencia

Recibido Dic. 26, 2018; Aceptado Feb. 28, 2019; Versión final Mar. 17, 2019, Publicado Oct. 2019

Resumen

El principal objetivo de este trabajo es responder si la academia chilena está o no preparada para formar especialistas en Big Data. Además de describir componentes teóricos y prácticos de Big Data junto con introducir una herramienta básica del tema, este trabajo define y presenta los resultados de una encuesta para explorar y analizar la realidad de la academia en Chile respecto al grado de viabilidad para capacitar y formar profesionales competentes en Big Data. **Con respecto a estos resultados, este trabajo argumenta que las condiciones necesarias para el desarrollo de competencias de Big Data en la academia chilena aún no están del todo presentes**, y así, la academia chilena necesita la adopción de tópicos y soluciones de Big Data para la formación futura de profesionales competentes de ese tan actual tema.

Palabras clave: Big Data; tecnología; información; grandes volúmenes de información; academia

Study of the Reality and Viability of the Education in Big Data in the Chilean Academy

Abstract

The main objective of this work is to answer whether the Chilean academy is or it is not prepared to educate Big Data specialists. Besides of describing theoretical and practical Big Data components along with introducing a basic tool of the subject, this work defines and presents the results of a survey to explore and analyze the reality of the academy in Chile regarding the degree of viability to educate and train competent professionals in Big Data. Concerning these results, this work argues that the necessary conditions for the development of Big Data competencies in the Chilean academy are not yet fully present, and thus, the Chilean academy needs to adopt Big Data topics and solutions for the future training of competent professionals of that current topic.

Keywords: Big Data; technology; information; large volumes of information; academy

INTRODUCCIÓN

Un dato es una representación simbólica de alguna propiedad observable de un objeto, evento, y/o de su ambiente el que sin una contextualización y propósito es poco relevante para la toma de decisiones (Baskarada y Koronios, 2013). Según (Chaffey y Wood, 2005), la información se corresponde con datos los que agregan valor para la comprensión de un tema en particular, de gran relevancia para las organizaciones en procesos de toma de decisiones y en la generación de conocimiento. Según Liew (2007), el conocimiento es un cuerpo de información para crear y aumentar el valor para la empresa. Tal como lo muestra la figura 1 (Davenport y Prusak, 2000), el procesamiento de los datos origina información, y el procesamiento de información genera conocimiento. Las pirámides DIKW (Data: datos, Information: información, Knowledge: conocimiento, Wisdom: sabiduría) (Liew, 2007) y DIKIW (Data: datos, Information: información, Knowledge: conocimiento, Intelligence: inteligencia, Wisdom: sabiduría) (Liew, 2013; Baskarada y Koronios, 2013) consideran la sabiduría y la inteligencia como resultados del procesamiento del conocimiento, y la sabiduría como resultado del procesamiento de la inteligencia, respectivamente. La base de todos estos contextos son los datos y su procesamiento para conseguir información.

Según White (2015), las bases de datos que usan lenguaje estructurado de consulta tipo SQL (del inglés Structured Query Language) trabajan sobre información estructurada, área de conocimiento y formación base en carreras técnicas y profesionales de las áreas de informática, ciencias de la computación, y de informática con negocios durante años en Chile. Actualmente, la información tiene cada vez menos estructura (Gandomi y Haider, 2015), y su volumen se acrecienta sobre los límites de sistemas computacionales tradicionales, lo que requiere una consideración en las bases formativas de las carreras de las áreas mencionadas. Por ejemplo, los sistemas computacionales en internet tales como Facebook y servidores de alojamiento web (hosting) trabajan con datos semi-estructurados y no estructurados, usualmente utilizados para la generación de información. Hoy se vive en la sociedad del Big Data, esto es, en una sociedad que requiere trabajar con grandes volúmenes de datos e información. Tal como lo argumenta White (2015), la cantidad de datos generados por máquinas, como parte de la Internet de las Cosas, llega a ser aún más grande que los datos que los seres humanos producen. Por ejemplo, los registros de autenticación y acceso (logs), los lectores de radiofrecuencia, los sensores de red, los sensores de GPS en vehículos, los registros de transacciones web contribuyen día a día de manera significativa al crecimiento del Big Data. La buena noticia es que el Big Data está aquí, pero aún no se desarrollan en Chile de forma masiva las competencias técnicas para su almacenamiento ni de conocimiento para su análisis.

En la educación superior en Chile y Latinoamérica, salvo trabajos de investigación tales como (Camargo et al., 2015; Valenzuela et al., 2016; Vidal et al., 2018), no existe una clara evidencia o tendencia para el desarrollo de competencias de dominio de Big Data. La formación de Ingenieros en Chile aun no incluye explícitamente cursos de Big Data (salvo que el tópico de Big Data fuera así en denominados optativos o talleres de especialización). El objetivo de este trabajo es presentar y describir la tendencia actual del procesamiento de grandes volúmenes de datos e información junto con algunas herramientas y metodologías de computación para dicho propósito, y así dar a conocer los detalles de un estudio exploratorio y los resultados de una encuesta acerca de si la academia en Chile está preparada para formar ingenieros especialistas en Big Data y si es factible su incorporación en la educación superior chilena.

Los sistemas de computación tradicional, software y hardware, presentan problemas para el trabajar con Big Data que se caracteriza por su volumen de datos, variedad de datos, y velocidad de cambio de datos. Si bien es cierto, siempre se intenta obtener un mayor soporte computacional en términos de hardware, los sistemas paralelos y distribuidos de computación son las soluciones, aun cuando los primeros representan un alto costo en dinero para su implementación, mientras que para los segundos, por las diferencias de espacio respecto de los primeros principalmente, además de no existir un gran respaldo software para la implementación de sistemas de computación.

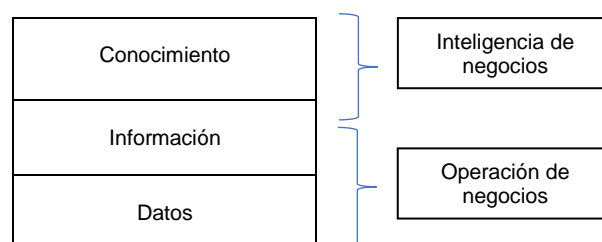


Fig. 1: Datos, información, conocimiento y su valor para las organizaciones.

En la historia de la computación, desde hace tiempo Google se caracteriza por la calidad y eficiencia de sus resultados de búsqueda. Existía una interrogante del sistema computacional software y hardware de Google, hasta que Dean y Ghemawat (2004) presentan detalles de la plataforma MapReduce. Desde entonces, el mundo del software libre trabaja en productos que logren el rendimiento de los productos de Google. A continuación, se describen la plataforma MapReduce de Google como base de la herramienta libre Hadoop, y la plataforma Preguel de Google que representa la base de herramientas libres tales como Giraph (Shaposhnik et al., 2015; Valenzuela et al., 2016; Vidal et al., 2018) y Spark (Guller, 2015; Shi et al., 2016). Tal y como lo describe el trabajo de Vidal et al. (2018), MapReduce representa una metodología de programación original de Google (Dean y Ghemawat, 2004; Dean y Ghemawat, 2010) para la computación distribuida sobre grandes volúmenes de datos o Big Data, con una amplia difusión y adopción de su implementación libre de código abierto Hadoop (Apache Hadoop, 2018; White, 2015; Lin y Dyer, 2013).

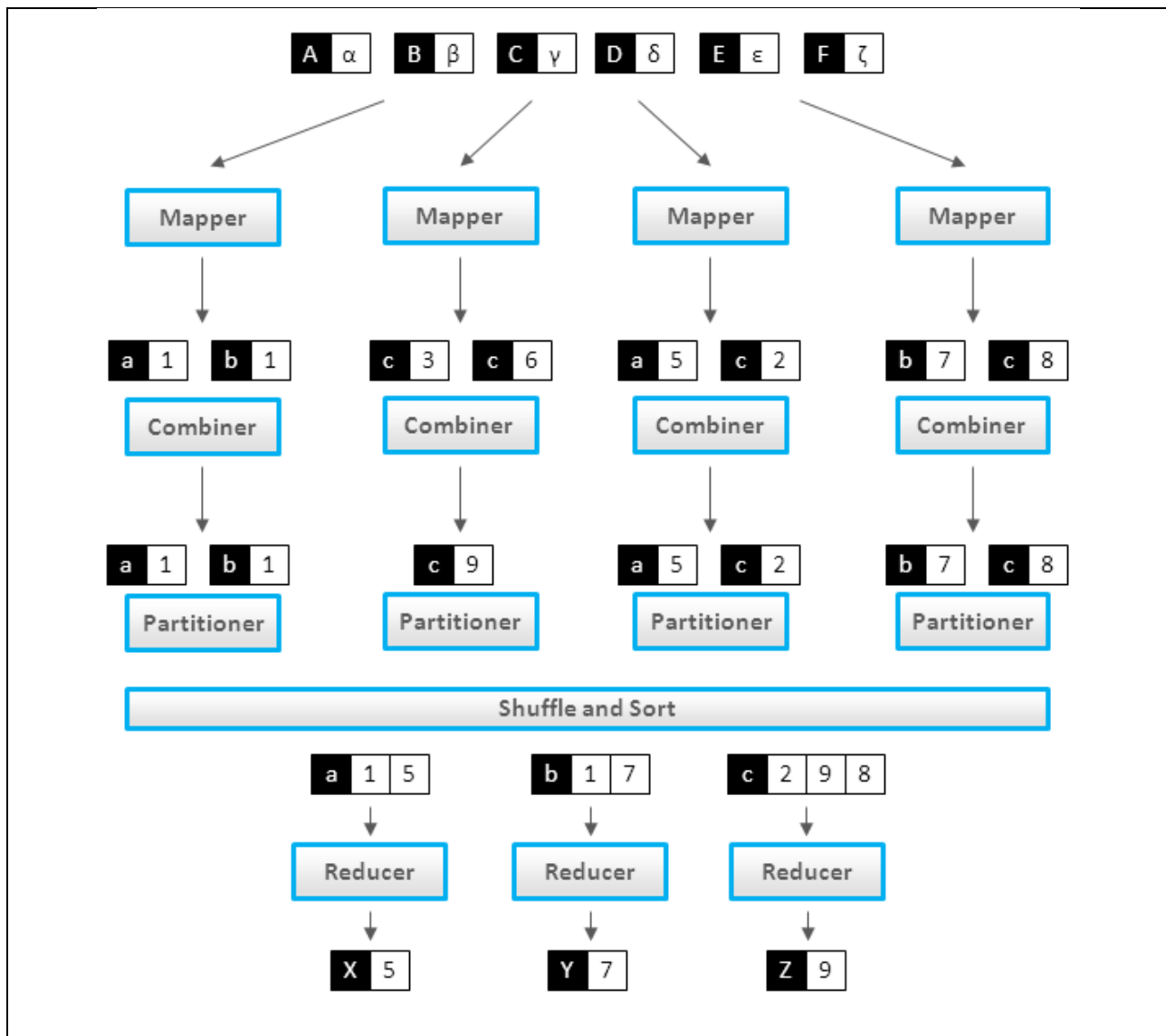


Fig. 2: Arquitectura de Funcionamiento de MapReduce.

La filosofía MapReduce se basa en la técnica algorítmica dividir-y-conquistar (Cormen et al., 2009; White, 2015; Hernández y Hernández, 2015) que consiste en, de manera recursiva, dividir un problema complejo en sub-problemas más pequeños para su solución los cuales pueden ser tratados de manera independiente cuando alcanzan un tamaño base. Así, MapReduce como sistema distribuido, divide los datos en porciones más pequeñas para su procesamiento en paralelo, esto es, o bien en paralelo por múltiples hilos trabajadores en un procesador, por múltiples procesadores trabajadores en una máquina multiprocesador, o por múltiples máquinas trabajadoras en un clúster o red de máquinas (White, 2015). La salida final de este proceso es la combinación de resultados intermedios de cada uno de sus trabajadores (mapeadores del inglés mappers, y reducers del inglés reducers). En la práctica, MapReduce es un marco de trabajo de computación distribuida que permite obviar costos de programación distribuida tales como garantizar el envío y recepción de mensajes.

La figura 2 (Vidal et al., 2018) ilustra la arquitectura de funcionamiento de la filosofía MapReduce de Hadoop: un programa de solución o trabajo en Hadoop ejecuta 4 pasos principales: i) División de datos donde múltiples fracciones de datos son entregadas a cada uno de los mapeadores, ii) Map donde se ejecutan las funciones Map para procesar los datos y procesador con la identificación de elementos relevantes y enviarlos a la etapa de organización, iii) Organización para organizar los datos y agrupar resultados intermedias, y entonces distribuirlos hacia la etapa de reducción, y iv) Reduce donde se ejecutan las funciones Reduce para compactar y resumir los resultados a ser escritos en disco (Galindo et al., 2016). Así, en Hadoop y MapReduce, objetos mapeadores ejecutan funciones map, y objetos reductores ejecutan funciones reduce. Para mayor detalle del funcionamiento y filosofía MapReduce se sugiere revisar (White, 2015; Vidal et al., 2018).

METODOLOGÍA

La tabla 1 detalla una encuesta aplicada a docentes pertenecientes al área de tecnologías e informática de algunas instituciones de educación superior en Chile. Específicamente, esta encuesta fue aplicada a 15 docentes pertenecientes a diferentes entidades chilenas de educación superior. La encuesta fue realizada a finales de 2016 y comienzos de 2017, y está compuesta de 15 preguntas clasificadas en 4 segmentos. Este estudio fue realizado mediante la plataforma de Google Docs bajo el nombre de Estado Actual del Big Data en la Academia Chilena (EstudioBigData, 2018). En las figuras 3 y 4 se muestra parte de la interfaz gráfica de la herramienta de este estudio así como un gráfico de las respuestas al primer conjunto de preguntas acerca del origen de los encuestados, respectivamente.

Tabla 1: Encuesta para docentes del área tecnología e informática de entidades de educación superior en Chile.

Número	Segmento	Pregunta	Tipo de Respuesta
1	Datos personales Encuestado	1 - 5	Opción única
2	Uso de habilidades, tecnología informática y de Big Data	6 - 8	Opción múltiple
3	Grado de Especialización en Tecnología informática y de Big Data	9 - 12	Opción única
4	Nivel de relevancia del Big Data para vuestra institución	13 -15	Opción única

Tabla 2: Preguntas acerca de Uso de habilidades, tecnología informática y de Big Data.

Pregunta	<i>En la siguiente tabla, seleccione aquellas técnicas, tecnologías o herramientas que alguna vez ha utilizado su institución</i>
Opciones Inclusivas	Base de datos (BD) NoSQL – Base de datos SQL – Hadoop – MapReduce – R – SAS – Machine Learning – Java – Hive QL – Python – Mahout – Tableau – Julia – Ipython – Ruby – Qlikview – Otra (indicar)
Pregunta	En la siguiente tabla, seleccione aquellas habilidades estadísticas que su institución, alguna vez ha utilizado
Opciones Inclusivas	Metodología para el procesamiento de Big Data – Estándares para el procesamiento de Big Data – Uso de software estadístico tales como: Excel, SAS, SPASS u otro similar – Habilidades de administración de datos incluyendo documentación, registro y control de acceso – Habilidad para trabajar con análisis de texto – Minería de Datos – Otra (indicar)
Pregunta	En la siguiente tabla, seleccione aquellas otras habilidades desarrolladas por su institución
Opciones Inclusivas	Comunicación – Creatividad para la solución de problemas – Trabajo en Equipo – Iniciativa – Privacidad – Gobierno de datos – Ética – Otra (indicar)

El estudio Estado Actual del Big Data en la Academia Chilena, como su nombre lo indica, busca medir y conocer la realidad actual del Big Data en la educación superior chilena, tal y como se aprecia en la tabla 1, mediante la cuantificación de las siguientes variables: uso previo de tecnología informáticas y de Big Data, nivel de especialización en tecnologías informáticas y de Big Data, nivel de relevancia del Big Data para vuestra institución académica. Las tablas 2, 3 y 4 presentan las preguntas asociadas a las variables de este estudio junto con sus respuestas usando una escala tipo Likert (Likert, 1932) para medir la viabilidad para la formación en Big Data en la academia chilena.

RESULTADOS Y DISCUSIÓN

Según los resultados de la encuesta, un 93% de los encuestados representa a docentes universitarios, y el 7% restante pertenece a otro tipo de institución de educación superior como se aprecia en la figura 4. En este contexto, tal y como se aprecia en la figura 5, un 40% de los encuestados cumple el rol de sólo docente, quienes no se dedican a investigar, una semi-limitante para con la adquisición y desarrollo de nuevas temáticas. Así mismo, un 40% de los docentes encuestados cumple un rol de docente / investigador quienes, en teoría, están preparados para investigar nuevas tecnologías, técnicas o metodologías tales como Big Data,

mientras un 20% de los encuestados señalan que su rol se refiere a otro, el que podría ser sólo de coordinador, director, o roles similares. Un 0% de los encuestados cumple un rol de sólo investigación.


ESTADO ACTUAL DEL BIG DATA EN LA ACADEMIA CHILENA

INTRODUCCIÓN

La presente encuesta tiene como objetivo evidenciar el estado actual del tópico "Big Data" en la academia Chilena

INSTRUCCIONES

Vuestra experiencia y visión como institución educacional que forma a futuros especialistas en un contexto de computación e informática es muy importante para este estudio. Sus respuestas serán tratadas de manera confidencial y no hay respuestas buenas ni malas



NOMBRE DE LA INSTITUCIÓN

(Tu respuesta)

DIRECCIÓN DE LA INSTITUCIÓN

(Tu respuesta)

NEXT

Página 1 de 3

Fig. 3: Interfaz gráfica de plataforma Google Docs para la realización de estudio Estado Actual del Big Data en la Academia Chilena.

Las figuras 6 y 7 muestran detalles de la productividad académica en el tópico de Big Data de las instituciones de educación superior de los encuestados, esto es, publicaciones en revistas indexadas o no indexadas, presentación de trabajos en congresos académicos o algún otro tipo de evento externo acerca del tema de Big Data, lo que corresponde a los resultados de las últimas preguntas de la tabla 4. Según estos resultados, la productividad académica de los encuestados con respecto al tópico de Big Data es muy baja (prácticamente cero), es decir, existe una clara correlación entre el rol mayoritario de los encuestados y la productividad académica de los mismos: dada la no dedicación exclusiva a investigación, poco es el trabajo de actualización de nuevas tecnologías como Big Data. Estos resultados son consecuentes con lo que muestra la figura 8 donde sólo un 13% de las instituciones de educación de los encuestados considera de alta relevancia el tópico de Big Data.

Cuando se preguntó acerca de las herramientas o técnicas utilizadas por la institución referente al tratamiento de datos e información (1era pregunta de tabla 3), en el ámbito orientado totalmente a Big Data, un 33% de los encuestados respondió que la más utilizada eran Bases de Datos NoSQL, aun cuando la adopción de una filosofía de como MapReduce es clave para comprender contextos, tipos de problemas y resultados posibles de solucionar mediante el trabajo con grandes volúmenes de datos. Además, un 73% de los encuestados respondió que la herramienta no relacionada con Big Data más usada corresponde a bases de datos SQL aunque dicho tipo de herramienta permite principalmente trabajar con volúmenes de información estructurada y relacional no del tamaño propio de Big Data.

Tabla 3: Preguntas acerca del Grado de Especialización en Tecnología informática y de Big Data.

Pregunta	<i>De acuerdo con el criterio señalado, indique el grado de especialización que usted tiene con las siguientes herramientas computacionales que se listan a continuación. Si usted no ha utilizado alguna de estas herramientas, se asume que el grado es "Nada" de conocimiento</i>
Ítems	Base de datos NoSQL – Base de datos SQL – Hadoop – MapReduce – R – SAS – Machine Learning – Java – Hive QL – Python – Mahout – Tableau – Julia – Ipython – Ruby – Qlikview – Otra (indicar)
Opciones Exclusivas para cada ítem.	Nada – Bajo – Medio – Avanzado – Experto
Pregunta	<i>De acuerdo con el criterio señalado, indique el grado de especialización que usted tiene con las siguientes habilidades estadísticas que se listan a continuación. Si usted no posee alguna de estas habilidades, se asume que el grado es "Nada" de conocimiento</i>
Ítems	Metodología para el procesamiento de Big Data – Estándares para el procesamiento de Big Data – Uso de software estadístico tales como Excel, SAS, SPSS u otro similar – Habilidades de Administración de datos incluyendo documentación, registro y control de acceso – Habilidad para trabajar con análisis de datos – Minería de datos – Otro (indicar)
Opciones Exclusivas para cada ítem.	Nada – Bajo – Medio – Avanzado – Experto
Pregunta	<i>De acuerdo con el criterio señalado, indique el grado de especialización que usted tiene con las siguientes habilidades que se listan a continuación. Si usted no posee alguna de estas habilidades, se asume que el grado es "Nada" de conocimiento</i>
Ítems	Comunicación – Creatividad para la solución de problemas – Trabajo en equipo – Iniciativa – Privacidad – Gobierno de Datos – Ética – Otro (indicar)
Opciones Exclusivas para cada ítem.	Nada – Bajo – Medio – Avanzado – Experto

Tabla 4: Preguntas acerca de Uso de habilidades, tecnología informática y de Big Data.

Pregunta	<i>¿Qué grado de importancia se consideran las técnicas y herramientas de BIG DATA en los planes y programas de vuestras carreras? (Seleccione su alternativa)</i>
Opciones Exclusivas	Nada – Bajo – Medio – Alto – Avanzado
Pregunta	<i>Número de trabajos de aplicación y/o investigación acerca de BIG DATA que se b realizado en vuestra institución (marque según cantidad)</i>
Opciones Exclusivas	0 – 1 – entre 1 y 3 – entre 3 y 5 – más de 5
Pregunta	<i>Número de eventos académicos en los cuales el tópico BIG DATA ha estado presente, en los que Ud. haya participado (marque según cantidad)</i>
Opciones Exclusivas	0 – 1 – entre 1 y 3 – entre 3 y 5 – más de 5

Claramente, por el uso masivo de la tecnología y la presencia de grandes volúmenes de datos (Big Data), las competencias relativas a Big Data son esenciales en diversos ámbitos actuales tales como empresarial y educación (Huda et al., 2018; Dede, 2016). Tal y como menciona Dede (2016), la investigación en educación se puede beneficiar enormemente del Big Data y la actual revolución informática más aun ahora que la disponibilidad de recursos en línea relacionados a tecnología educativa amplía las oportunidades de investigación en el aprendizaje, incluidas las relacionadas con el género, el origen étnico, la situación económica, entre otras variables. La educación y la investigación en Chile deberían considerar estos recursos para el proceso de enseñanza / aprendizaje para el desarrollo de competencias de Big Data.

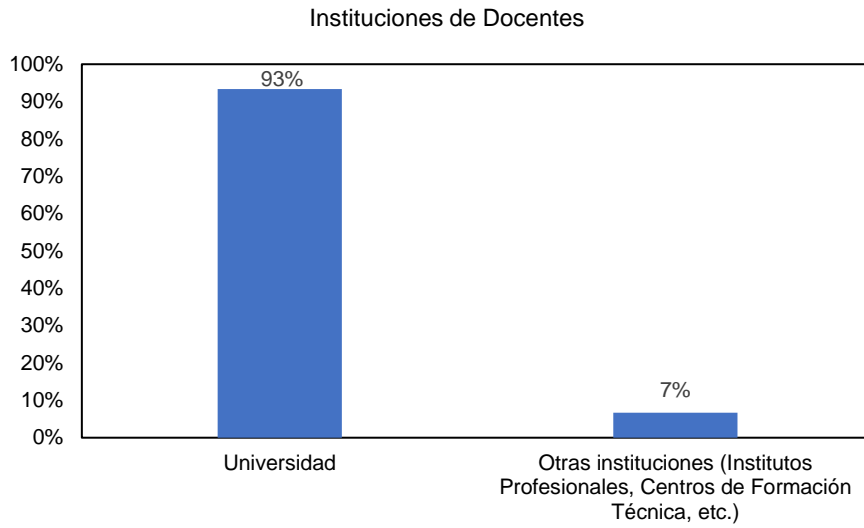


Fig. 4: Instituciones de educación superior de los encuestados.

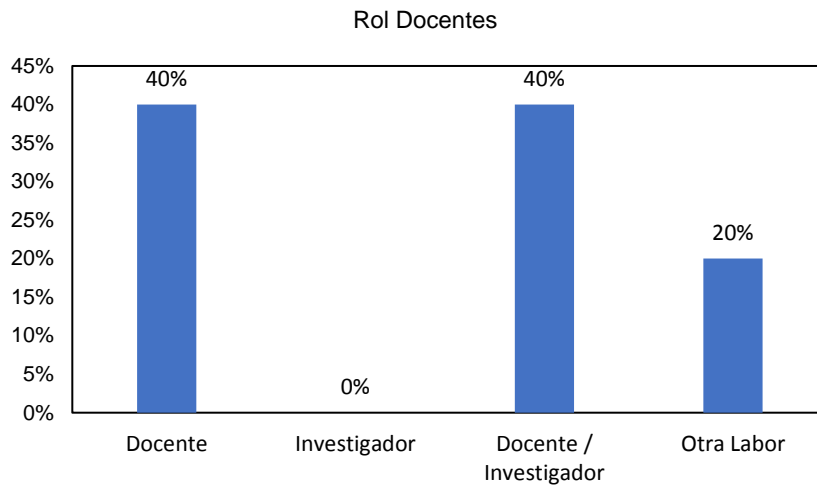


Fig. 5: Rol principal de los encuestados.

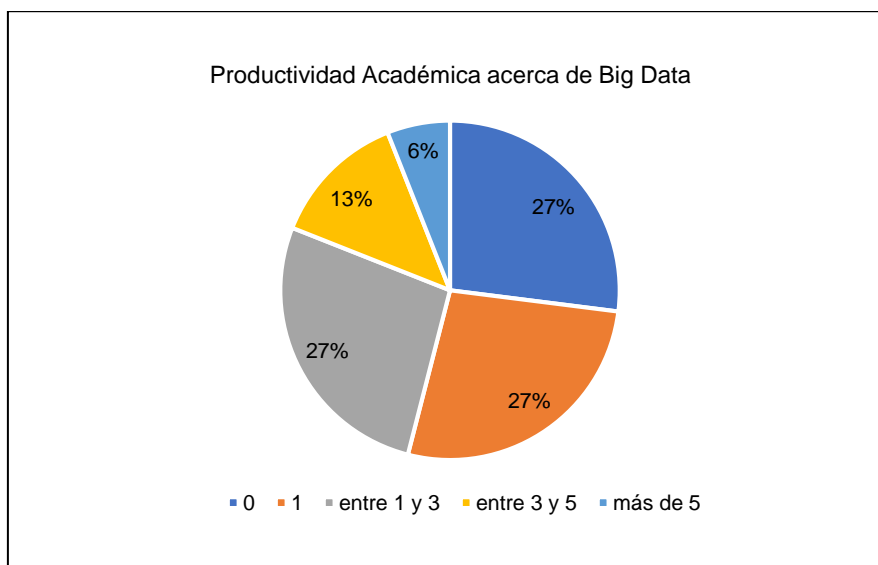


Fig. 6: Productividad académica de las instituciones de los entrevistados en el tópico de Big Data.

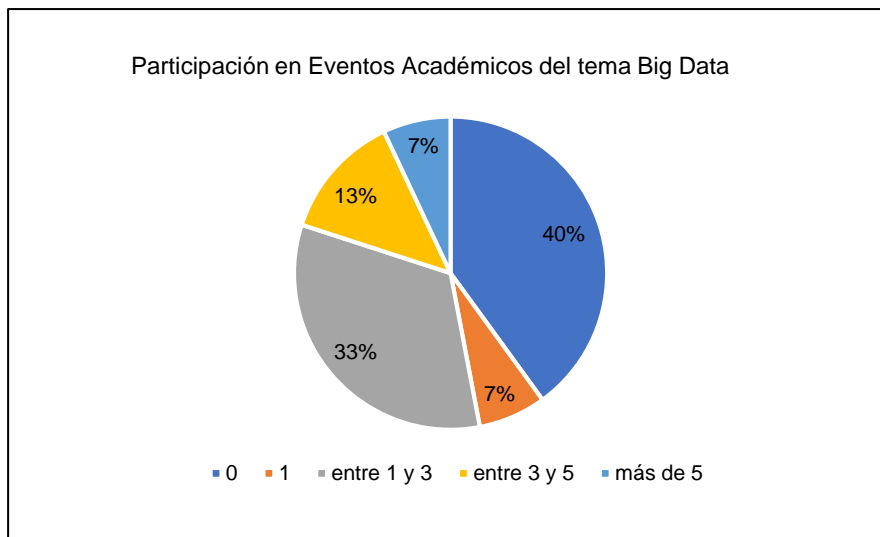


Fig. 7: Participación de los entrevistados en eventos académicos del tópico de Big Data.

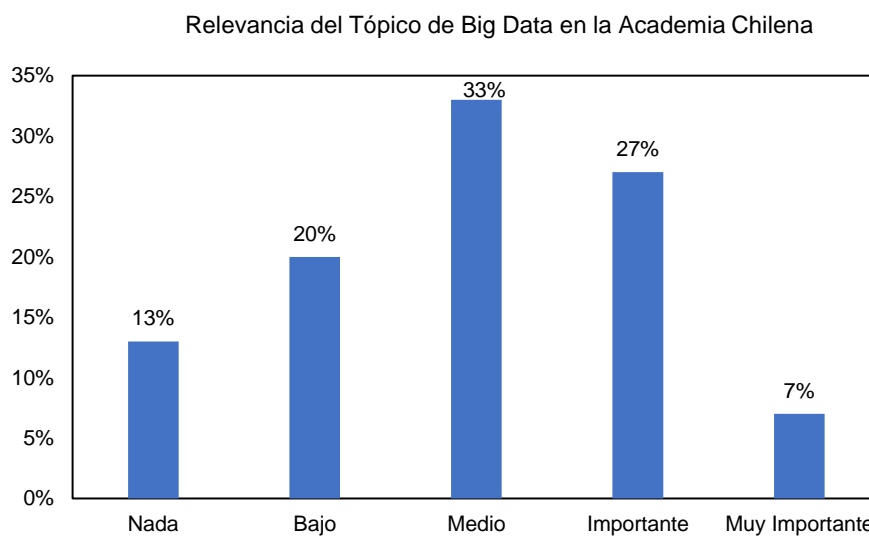


Fig. 8: Relevancia actual del tópico Big Data para instituciones de educación superior en Chile.

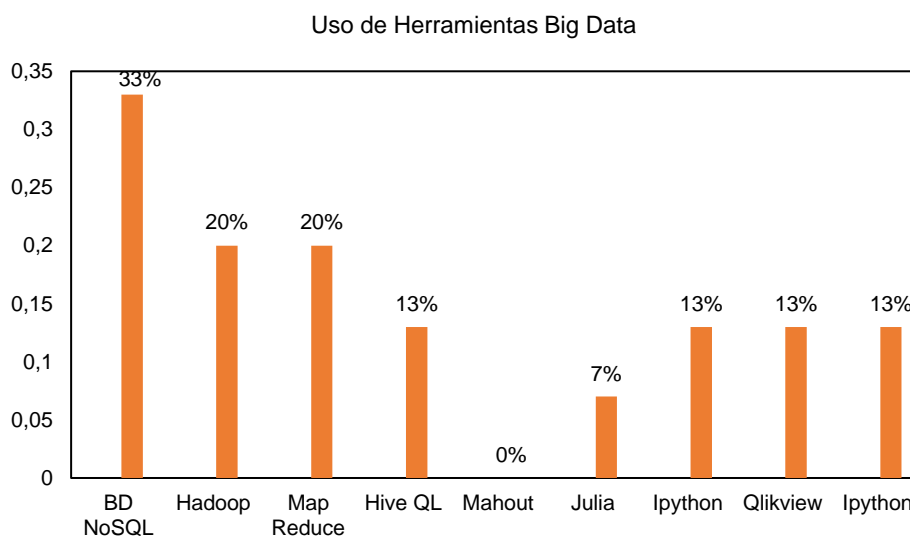


Fig. 9: Porcentajes de uso y aplicación de herramientas Big Data en instituciones de educación en Chile.

De acuerdo con el trabajo de Williams (2016), la analítica del aprendizaje (del inglés analytical learning) y el uso masivo de información o Big Data juegan un rol preponderantemente positivo para el desarrollo de competencias en estudiantes universitarios. Tal y como se señaló previamente, esto invita a las universidades en Chile al uso guiado de las plataformas de aprendizaje existentes como herramientas de apoyo a cursos de temáticas de Big Data (y también de cualquier otra competencia).

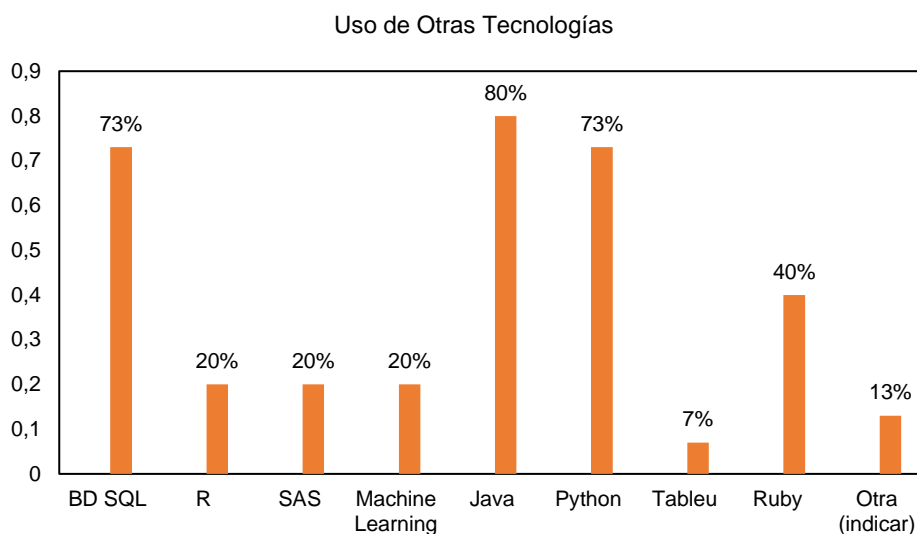


Fig. 10: Porcentajes de uso y aplicación de herramientas de informática y computación no Big Data en instituciones de educación en Chile.

DISCUSIÓN Y ANÁLISIS FINAL

La incorporación del Big Data en la academia requiere la inclusión de tópicos de análisis, diseño, implementación y prueba de soluciones para el área informática y computación, y de tópicos de tratamiento de información y datos tanto para informática como para cualquier otra área que trabaje con información, es decir, en la actualidad para cualquier área de conocimiento. Este trabajo ha presentado los resultados de un estudio acerca de la realidad de Big Data en la academia chilena del área informática. Según estos resultados, en la academia en Chile, aun cuando existen las bases para ingresar al mundo del Big Data, se requiere un cambio estructural en contenidos teóricos y prácticos para lograr una formación profesional competente en Big Data. Además, la realización efectiva de dicho cambio necesita una definición de competencias y resultados de aprendizaje, así como niveles de desarrollo. Incluir Big Data en la academia informática en Chile está siendo similar a la incorporación del desarrollo de software orientado a objetos. Por ejemplo, uno de autores no desarrolló competencias de orientación a objetos en el pregrado (1998-2003), y dicho paradigma existía en la industria y academia mundial desde fines de los años 60 (Dahl et al., 1968) y mediados de los 80, respectivamente. Big Data es conocido gracias a Google desde 2004 (Dean y Ghemawat, 2004), y, con el tiempo transcurrido, este trabajo muestra la realidad de la academia en Chile en la viabilidad de formar profesionales competentes en Big Data.

CONCLUSIONES

Este trabajo ha mostrado la definición y análisis de resultados de un estudio acerca de la viabilidad de formar profesionales de las áreas de informática y computación competentes en el tema de Big Data en la academia en Chile. Según los resultados y análisis de estos, este trabajo argumentó que aún no están todas las condiciones presentes para la formación de profesionales informáticos especialistas en Big Data, a pesar de la masiva relevancia y difusión del tema. Así, este trabajo planteó que la inclusión de tópicos de Big Data es necesaria para la formación de las nuevas generaciones de ingenieros de las áreas de informática y computación en Chile, ya que actualmente la demanda de profesionales especialistas en dicho tópico es cada vez mayor, y el dominio de estos tópicos son absolutamente necesarios para el desarrollo de soluciones a problemáticas actuales de Big Data. Ya están las bases de programación, modelamiento, e ingeniería de software necesarias para una adopción exitosa de Big Data en la academia chilena, y sólo basta coordinar su inclusión para la formación de profesionales e investigadores de tópicos de Big Data.

REFERENCIAS

- Baskarada, S. y A. Koronios, Data, Information, Knowledge, Wisdom (DIKW): A Semiotic Theoretical and Empirical Exploration of the Hierarchy and its Quality Dimension, doi: 10.3127/ajis.v18i1.748, Australasian Journal of Information Systems, 18(1), Noviembre (2013)
- Camargo, J. J., J. F. Camargo y L. Joyanes, Conociendo Big Data, Revista Facultad de Ingeniería, 24(38), 63-77, Enero-Abril (2015)
- Chaffey, D. y S. Wood, Business Information Management: Improving Performance Using Information Systems, Harlow, FT Prentice Hall (2005)
- Dahl, O.J., B. Myhrhaug y K. Nygaard, Some features of the SIMULA 67 language, Proceedings of the Second Conference on Applications of Simulations, 29 – 31, Diciembre (1968)
- Davenport, T. y D. D. Prusak, Working Knowledge: How Organizations Manage What They Know, doi: 10.1145/347634.348775, Journal Ubiquity, ACM, USA, Agosto (2000)
- Dean J. y S. Ghemawat, MapReduce: a Flexible Data Processing Tool, doi: 10.1145/1629175.1629198, Communications of the ACM, 53(1), 72-77 (2010)
- Dean J. y S. Ghemawat, MapReduce: Simplified Data Processing on Large Clusters, en Actas de 6th Conference on Symposium on Operating Systems Design & Implementation OSDI 2004, San Francisco, CA, USA, 137-150 (2004)
- Dede, C., Next steps for “Big Data” in education: Utilizing data-intensive research, Educational Technology Harvard, LVI (2), 37 – 42 (2016)
- EstudioBigData, Estado Actual del Big Data en la Academia Chilena, Universidad Bernardo O’Higgins (2018)
- Galindo J., M. Acher y otros cuatro autores, Exploiting the enumeration of All Feature Model Configurations - A New Perspective with Distributed Computing, doi:10.1145/2934466.2934478, en Actas de SPLC 2016, Beijing, China, Septiembre (2016)
- Gandomi, A. y M. Haider, Beyond the hype, doi: 10.1016/j.ijinfomgt.2014.10.007, International Journal of Information Management, 35(2), 137–144, Abril (2015)
- Guller, M., Big Data Analytics with Spark., ISBN: 978-1-4842-0965-3, Apress – Spring, New York, NY, USA (2015)
- Hernández, A. y A. Hernández, Acerca de la aplicación de MapReduce + Hadoop en el tratamiento de Big Data, Revista Cubana de Ciencias Informáticas, 9(3), 49-62, Julio-Septiembre (2015)
- Huda, M., A. Maselena y otros cinco autores, Big Data Emerging Technology: Insights into Innovative Environment for Online Learning Resources, doi: 10.3991/ijet.v13i01.6990, International Journal of Emerging Technologies in Learning (IJET), 13(01), Enero (2018)
- Liew, A., Understanding Data, Information, Knowledge and their interrelationships, journal of Knowledge Management Practice, 7, Junio (2007)
- Liew, A., DIKIW: Data, Information, Knowledge, Intelligence, Wisdom and their Interrelationships, journal Business Management Dynamics, 10(2), 49-62, Abril (2013)
- Likert, R., A Technique for the Measurement of Attitudes, Archives of Psychology, 140, 1-55 (1932)
- Lin J. y C. Dyer, Data-Intensive Text Processing with MapReduce, ISBN: 1608453421, Morgan and Claypool Publishers, USA (2010)
- Shaposhnik, R., C. Martella y D. Logothetis, Practical Graph Analytics with Apache Giraph, ISBN: 1484212525, Apress, USA (2015)
- Shi, J., Y. Qiu y otros cinco autores, Clash of the Titan: MapReduce vs. Spark for Large Scale Data Analytics, doi: 10.14778/2831360.2831365, en Actas de Very Large Data Bases (VLDB), September 5th-09th, New Delhi, India (2016)
- Valenzuela, S., C. Vidal, J. Morales y L. López, Ejemplos de Aplicabilidad de Giraph y Hadoop para el Procesamiento de Grandes Grafos, doi: 10.4067/S0718-07642016000500019, Información Tecnológica, 27 (5), 171-180 (2016)
- Vidal, C., M. Bustamante, M. Lappo y M. Núñez, En la Búsqueda de Soluciones MapReduce Modulares para el Trabajo con BigData: Hadoop Orientado a Aspectos, doi: 10.4067/S0718-07642018000200133, Información Tecnológica, 29 (2), 133-140 (2018)
- White, T., Hadoop: The Definitive Guide, ISBN: 1491901632, O’Reilly, 4th Ed., Sebastopol, CA, USA (2015)
- Williams, P., Assessing collaborative learning: big data, analytics and university futures: big data, analytics and university futures, doi: 10.1080/02602938.2016.1216084, Assessment & Evaluation in Higher Education, Julio (2016)