



EXPECTATIVAS EN TORNO A BIG DATA

La voz de profesores y estudiantes de educación superior

EXPECTATIONS OVER BIG DATA
The voice of higher education's professors and students

JULIO LÓPEZ-NÚÑEZ¹, GERARDO CERDA-NEUMANN¹
¹ Universidad Gabriela Mistral, Chile

KEYWORDS

Data Processing
Big Data
Higher Education
Vocational Training
Vocational Guidance

ABSTRACT

Data has become a significant element in understanding the behavior of various systems. In this context, Big Data is a concept that refers to working with large volumes of data, with a blurred boundary between data analysis and data management. To this end, this research is framed in the qualitative paradigm through interviews with students and teachers participating in an introductory course on Big Data. It seeks to unveil the utility surrounding this topic. The main findings show a divergence between Big Data and what is understood of him.

PALABRAS CLAVE

Procesamiento de datos
Big Data
Educación Superior
Formación Profesional
Orientación profesional

RESUMEN

Los datos se han convertido en un elemento significativo para comprender el comportamiento de diversos sistemas. En este contexto, Big Data es un concepto que hace referencia al trabajo con grandes volúmenes de datos y con una frontera difusa entre análisis y gestión de ellos. Con tal objeto, esta investigación se enmarca en el paradigma cualitativo. Mediante entrevistas a estudiantes y profesores, partícipes de un curso introductorio a Big Data, busca develar la utilidad en torno a esta temática. Los principales hallazgos constatan una divergencia entre el concepto y lo que se entiende de él.

Recibido: 19/06/2022
Aceptado: 22/08/2022

1. Introducción

En la literatura científica no es sencillo encontrar una definición absolutamente aceptada respecto al concepto de *Big Data*. En este contexto, Joyanes (2013) realiza una aproximación al concepto al citar, en su libro titulado *Big Data*, la definición propuesta por Gartner “*Big Data* excede el alcance de los entornos de hardware de uso común y herramientas de software para capturar, gestionar y procesar los datos dentro de un tiempo transcurrido tolerable para su población de usuarios”. Bajo esta perspectiva, *Big Data* surge como respuesta a la incapacidad de procesar un volumen de datos existente mediante las tecnologías tradicionales de que disponen las organizaciones. Es decir, *Big Data* es útil para presentar una solución a las limitaciones de los sistemas informáticos.

En este marco, cobra relevancia la definición del *McKinsey Global Institute*, organismo que centra su definición en las capacidades del software para procesar un conjunto de datos. En detalle, la definición se sustenta en las capacidades de procesamiento de un *DBMS*: “*Big Data* se refiere a los conjuntos de datos cuyo tamaño está más allá de las capacidades de las herramientas típicas de software de bases de datos para capturar, almacenar, gestionar y analizar” (Joyanes, 2013, p. 2).

En tanto, López-Núñez y Cerda-Neumann (2021) señalan que *Big Data* es “un concepto que permite describir un conjunto de datos que, dada su velocidad de generación, logra alcanzar una magnitud tal que solo es posible de ser procesado vía el uso de la computación distribuida” (p. 8). En este sentido, como una definición que trasciende en el tiempo, los autores citados ponen como centro de gravedad la arquitectura distribuida y las características de los datos (volumen y velocidad) a ser procesados. En términos prácticos, si los datos, independiente de su volumen, velocidad y variedad, pueden ser procesados por los sistemas informáticos habituales de la organización, se asume que dicho caso no es representativo de *Big Data*.

Señalado lo anterior, en una era informática marcada por el *Big Data*, es plausible reconocer que estos volúmenes de datos son el capital más importante de toda organización (Cukier y Mayer-Schönberger, 2013). Con tal objeto, diversas empresas descubrieron la posibilidad de potenciar la investigación de mercados vía el procesamiento de grandes conjuntos de datos (Cao, 2021).

Como suele suceder, las organizaciones, al vislumbrar los posibles beneficios del *Big Data*, aumentaron la apuesta y solicitaron el análisis continuo de los datos. Esta acción, dentro de muchos objetivos, se visualiza como una opción para extraer información y conocimiento a partir de datos (Cukier y Mayer-Schönberger, 2013).

Dado este contexto, los resultados obtenidos por la encuesta de intenciones de gasto en tecnología de la consultora Enterprise Strategy Group (2021), como también el estudio recientemente presentado por NewVantage Partners (2022), se ajustan a esta tendencia de valoración de los datos. En particular, ambos estudios muestran una alta apreciación de las empresas respecto a la implementación de proyecto *Big Data*. Específicamente, el 91,7% de las empresas consultadas señaló estar aumentando sus inversiones en proyectos de *Big Data* y otros datos, como también en iniciativas de inteligencia artificial. Frente a esta realidad, sin lugar a dudas podemos sostener la creciente necesidad de profesionales en esta materia.

Bajo esta línea argumentativa, desde la irrupción de los análisis sobre las tasas de retorno en educación (Psacharopoulos, 1973), propuesta emanada por el Banco Mundial, se observa una creciente alza en la educación meramente técnica (Lauglo, 1996; Bennell, 1996; Heyneman, 1997; Heyneman, 2003). Consecuentemente, es posible señalar que la educación superior se enfoca directamente hacia la producción de una mano de obra (Althusser, 1970; Bowles y Gintis, 1981; Espinoza, 2017) que satisfaga las necesidades crecientes del mundo laboral y, últimamente enfocada en un mercado globalizado (Maldonado, 2000; Bonal, 2002; Kempner y Loureiro, 2002).

En este orden de ideas, la incorporación de *Big Data* en el currículo de diversas carreras de educación superior se torna una necesidad (Sledgianowski, Goma y Tan, 2017; Rezaee y Wang, 2017). Enfatizando el punto anterior, durante los últimos años se observan variadas investigaciones que enfatizan la necesidad para experimentar y analizar *Big Data* en Instituciones de Educación Superior (Anshari. Alas y Guan, 2016; Dede, 2016; Hilbert, 2016; Long y Siemens, 2011; Waller y Fawcett, 2013). Sin duda, preparar profesionales para asumir la demanda del sector productivo, tanto local como global, es un desafío para dicha educación superior. Sin embargo, su complejidad se basa en diversas líneas, principalmente sobre las expectativas y la empleabilidad que tendría el *Big Data*.

2. Objetivos

Respecto a los objetivos se decidió utilizar la siguiente pregunta de investigación: ¿Es de utilidad para un estudiante de Ingeniería Informática (sistemas) participar en un curso de *Big Data*?

En este sentido, se debe señalar que el concepto valioso se entenderá como la acción que propende a facilitar al estudiante herramientas para desarrollar mejores soluciones informáticas y como consecuencia de eso, aumentar la empleabilidad.

Consecuente con lo anterior, para este trabajo se fijó como objeto de estudio la pertinencia de dictar un curso de *Big Data* en una carrera de Ingeniería Informática (sistemas). Ajustándose a la definición de López-Núñez

y Cerda-Neumann (2021), los conceptos teóricos involucrados en este trabajo son los grandes volúmenes de datos, procesamiento de datos y tecnología informática de procesamiento. Por último, los conceptos analíticos pertinentes a este trabajo se refieren al valor (aporte) del concepto *Big Data* para un futuro Ingeniero en Informática (sistemas).

3. Metodología

En cuanto al diseño metodológico que regula este trabajo, una primera definición logra situar la investigación dentro del paradigma cualitativo. Así mismo, la investigación es de carácter exploratorio, cuyo objetivo se centra en develar la utilidad para estudiantes y profesores respecto a un curso introductorio de *Big Data* desde su propio punto de vista y experiencia.

Respecto a la muestra, esta es de tipo intencional y opinática, logrando recolectar las vivencias y expectativas de 40 participantes (de los cuales el 10% corresponden a la categoría profesor). La muestra considera un rango de edad entre los 21 y los 46 y una mediana de 24 años para la categoría estudiante. Y, en el caso de los profesores, la media de edad es de 47 años. En cuanto al criterio de inclusión se exige ser partícipe de un curso introductorio al *Big Data*, junto a lo anterior, es requisito indispensable el deseo voluntario para participar en esta investigación.

Para la selección de los participantes se consideraron a estudiantes y profesores de dos instituciones de educación superior (IES), ambas con presencia a nivel nacional (Chile) y que se encuentran acreditadas ante la Comisión Nacional de Acreditación de Chile (CNA-Chile).

Respecto a la recolección de los relatos, considerando el contexto sanitario a propósito de la pandemia COVID-19, se optó por la aplicación de un formulario en formato electrónico, utilizando la plataforma Google Forms. En esta línea, y considerando los objetivos planteados para esta investigación, el formulario fue constituido por 11 reactivos de tipo ensayo y cuatro de tipo restringido, estos últimos enfocados en la caracterización de los participantes.

Con tal objeto, el instrumento centró su interés en develar la utilidad que tiene participar de un curso introductorio a *Big Data*, específicamente, en cuanto a los aprendizajes que se deberían alcanzar al finalizarlo. En esta línea, el instrumento aplicado consideró cinco reactivos enfocados en describir como se estructura el relato en torno a *Big Data*, y su distancia respecto a la definición aportada por Lopez-Nunez y Cerda-Neumann (2021). Además, el instrumento aplicado consideró seis reactivos centrados en develar las ideas circundantes a los aprendizajes esperados para una asignatura de este tipo. En tanto, las preguntas enfocadas a la caracterización de los entrevistados consideró tres ítems, cuya finalidad era conocer su edad, sus años de experiencia laboral y el área de las tecnologías de la información donde participan.

Respecto a la aplicación del instrumento, se consideró un contexto propicio para la temática central del estudio. Coincidente con esto, durante la primera semana de clases del semestre académico del año 2020, se optó por realizar esta actividad en el marco de una clase de *Big Data*. Así, los investigadores principales acudieron al salón y solicitaron la participación voluntaria de estudiantes y profesores. Para tal efecto, se destinaron en promedio 30 minutos para responder los 14 reactivos, los cuales podían ser accedidos a través de una plataforma web (<https://forms.gle/n6p5y8WYZo6kj24F7>).

Bajo este contexto, es imperioso señalar que la validación del instrumento fue realizada desde los aportes que entrega el método Delphi. Dada la complejidad que resulta de la exploración de esta temática, tanto Linstone y Turoff (1975), como también Okoli y Pawlowski (2004), proponen que esta técnica es efectiva para consensuar ideas. Esto facilita resolver un problema, desde los aportes que entregan los expertos en una determinada temática. Por consiguiente, el cuestionario aplicado a los participantes de esta investigación, es el fruto de dos rondas de evaluación del instrumento, cuyo objetivo fue develar la utilidad que tiene, para el alumnado y profesorado, un curso introductorio a *Big Data*. Particularmente, el grupo de expertos que trabajaron de forma anónima durante la evaluación, fue constituido por tres académicos, pertenecientes a IES que no formaron parte de los informantes clave de este trabajo.

Respecto al análisis de los datos, este fue de tipo temático, utilizando la descripción y/o interpretación del contenido de los mismos. Debido a esto, el trabajo consideró dos categorías de análisis, a saber: Conocimiento de *Big Data* y la Utilidad sobre el concepto. En esta línea, el análisis de los datos se realizó en cuatro etapas, comenzando con la a) Fase de preparación, cuyas actividades se centraron en descargar y preparar los datos almacenados en la herramienta *Google Forms*. Posteriormente se siguió con la b) Fase de pre análisis, ocasión en que los investigadores principales realizaron una lectura pormenorizada de las respuestas recolectadas, distribuyendo los relatos según las dos categorías de análisis predefinidas. Acto seguido se continuó con la c) Fase de análisis, instancia utilizada para codificar las citas representativas de cada relato y, mediante un proceso de progresiva abstracción, detectar sub temas mediante la agrupación de códigos con significados similares y/o relacionados entre sí. La última instancia corresponde a la d) Fase de relativización, ocasión propicia para la verificación y contraste de los hallazgos que surgen de los relatos de los participantes.

Con todo, resulta necesario destacar que este proceso de análisis es circular y reflexivo, yendo de lo general a lo específico y a la inversa para verificar la correcta interpretación de los datos. Por último, como resguardo ético se considera el consentimiento informado (Tójar y Serrano, 2000) de todos los participantes y, con tal de mantener la confidencialidad de los mismos, se optó por la utilización de códigos para referenciar a cada individuo que respondió el instrumento.

4. Resultados

Como ya fue señalado en el marco metodológico, las entrevistas realizadas consideraron la participación de estudiantes y profesores de un curso introductorio a *Big Data*, en esta línea se lograron definir dos categorías para el análisis. A saber: Conocimiento de *Big Data* y la Utilidad del tema.

Respecto al segmento de profesor, los resultados obtenidos desde la perspectiva del conocimiento y valoración de *Big Data*, es posible señalar que las respuestas son bastantes homogéneas. Así es como los profesores entrevistados, quienes poseen en promedio más de 20 años de experiencia en el área de las Tecnologías de Información (TI), manifiestan una alta valoración respecto a un curso introductorio a *Big Data*. Sin embargo, las razones de esta valoración varían. Para algunos es una herramienta necesaria, con tal de procesar los grandes volúmenes de datos, realidad patente en diversas organizaciones. También, los entrevistados señalan que *Big Data* es, sin lugar a dudas, muy necesario para apoyar la transformación digital.

“Todas las asignaturas de datos son claves actualmente para los profesionales ya que todas las empresas avanzan en esta línea como parte de la transformación digital”. (Profesor 1)

“En la actualidad es una herramienta muy utilizada para poder analizar la empresa”. (Profesor 2)

“Porque hoy en día las diversas actividades humanas están siendo muy impactadas por los volúmenes de datos que cada vez van creciendo y además se da el tema de tiempo real. Además, cada vez es más creciente la toma de decisiones asistida o autónoma en diferentes tipos de problemas”. (Profesor 3)

“Porque en el futuro cercano, una correcta gestión de grandes volúmenes de datos será un requisito base para los profesionales en general y en especial para aquellos ligados a las TIC”. (Profesor 4)

En tanto, respecto a la infraestructura necesaria, existe consenso en el uso de la nube (*Cloud Computing*) como plataforma de soporte para la aplicación de *Big Data*. En esta línea, ninguno de los entrevistados fue capaz de opinar respecto a la robustez de dicha tecnología. Esto se debe a la dificultad de medir esa robustez, fundamentado particularmente por el hecho que el volumen de datos a procesar crece diariamente.

“La idea es que sea en alguna nube, pero como PaaS y no tener que configurar todos los componentes”. (Profesor 1)

“La nube ayuda en el sentido que nuestros alumnos no siempre cuentan con el hardware necesario y ellos también valoran nuevas experiencias, en este caso trabajando en la nube con herramientas adecuadas”. (Profesor 3)

“Me parece que la infraestructura propuesta por Amazon EMR es una muy buena alternativa”. (Profesor 4)

Respecto a la principal competencia/habilidad/conocimiento que la asignatura *Big Data* debería entregar, hubo discrepancias notorias. A este respecto, los temas recurrentes en los discursos de los entrevistados se expresan en 1. La computación distribuida, 2. Análisis, lógica y habilidades comunicacionales, 3. Toma de decisiones y, 4. Modelamiento de soluciones usando la arquitectura distribuida. Esta disparidad de opiniones puede estar sustentada por la construcción sobre el término *Big Data* que cada profesor posee, en especial por el trabajo que desarrolla en la industria TI. Esto último se refuerza, toda vez que la principal competencia / habilidad / conocimiento que una empresa espera de un Ingeniero capacitado en *Big Data*, según los informantes clave, debería ser:

1. La ingesta de datos,
2. El análisis de datos,
3. Agregar valor mediante el procesamiento de datos de bajo costo y,
4. La capacidad de armar una arquitectura que permita la aplicación de Big Data.

“Yo diría que en el contexto de las funciones que ellos desarrollaran en empresas reales, el tema de la competitividad, toma de decisiones etc. implica trabajar conociendo, tener la actitud adecuada y sabiendo poner en práctica los conceptos, herramientas para trabajar con volumen, velocidad y variedad de datos estructuradas y no, relacionales o no que además se pueden trabajar en nubes con herramientas específicas,

de modo de agregar valor a los procesos de negocio donde sus empresas deban generar ventajas competitivas sostenibles”. (Profesor 3)

“La capacidad de poder modelar soluciones que consideren el almacenamiento y procesamiento de grandes volúmenes de datos en base a arquitecturas distribuidas robustas”. (Profesor 4)

Bajo esta perspectiva, la importancia de tener un curso introductorio a *Big Data*, los estudiantes entrevistados declaran una variación de respuestas que van desde medianamente importante a muy importante.

Para aquellos que consideran el máximo de importancia, esto se debe a que *Big Data* es visto como una línea de trabajo con mucho futuro. Su principal fortaleza está en ser considerada como una competencia laboral altamente demandada. Adicionalmente, algunos participantes manifiestan un interés de conocer *Big Data*, dado por la relevancia laboral percibida en el entorno. También es posible visualizar una percepción sobre *Big Data* como complemento de la Minería de Datos. Cabe indicar también que *Big Data* es considerado un aporte a la toma de decisiones y, por sobre todo, por la percepción que el mercado laboral así lo demanda.

“Principalmente porque está relacionado con el área de datos (Minería de Datos y Machine Learning), el cual además de ser unos de los ramos directamente relacionados con la carrera, me interesa bastante por la relación con el área y en donde este probablemente uno de los más usado afuera en la industria (no siempre, pero si en ciertos ámbitos)”. (Estudiante 1)

“Esta en conjunto con Minería nos permite trabajar con grandes cantidades de datos”. (Estudiante 2)

Para aquellos entrevistados que asignan un puntaje mediano alto a la relevancia de *Big Data* en su formación, se trata de un grupo de estudiantes que tienen poca o ninguna experiencia laboral. En tanto, aquellos estudiantes que asignan una baja valoración a la relevancia de la temática consultada, se debe particularmente a una falta de interés en el tema y la dificultad que perciben en los contenidos tratados en la asignatura.

“No tengo tanto interés, ya que de momento estoy centrado completamente en el desarrollo (programación)”. (Estudiante 4)

“Al ser muchos cursos los que debemos abarcar durante el semestre, como se trata de realizar trabajos grupales, entonces cada uno se especializó en un ramo”. (Estudiante 8)

Respecto a la infraestructura necesaria para *Big Data*, no existe consenso. Algunos simplemente declaran no entender la pregunta y otros indican que basta un computador “potente”, lo que claramente indica no entender el concepto de *Big Data*. Esto se comprende por su falta de conocimiento respecto a la imposibilidad de procesar los datos en un tiempo razonable, usando la tecnología tradicional. Algunos pocos mencionan arquitecturas específicas como AWS, Cloud, Hadoop o SaaS. En este caso, esas menciones son de los estudiantes que sí tienen experiencia laboral. Finalmente, ninguno de los entrevistados fue capaz de opinar respecto a la robustez de dicha tecnología. En consonancia a lo anterior, la principal competencia/habilidad/conocimiento que la asignatura *Big Data* entrega a su formación, los estudiantes manifestaron disímiles respuestas. Incluso, varios opinan no saber. En esta línea, las opiniones más relevantes apuntan hacia temáticas que dicen relación con:

1. Trabajo en equipo,
2. Análisis/Analítica, siendo la que más se repite.
3. Manejo de datos, siendo la segunda que más se repite y,
4. Minería de Datos.

“Saber ocupar Docker con Apache Hadoop y Scala”. (Estudiante 4)

“Conocimientos en Docker”. (Estudiante 12)

“Apostar por software libre: Hadoop y Spark”. (Estudiante 10)

“Ambari, con la máquina virtual de Cloudera (Hortonworks HDP)”. (Estudiante 12)

“Docker / Spark / Hadoop”. (Estudiante 9)

“Solo un PC que tenga la potencia necesaria”. (Estudiante 3)

Del mismo modo, la principal competencia/habilidad/conocimiento que una empresa espera de un Ingeniero capacitado en *Big Data*, también hubo un desacuerdo total. Cabe destacar, que la temática con mayor frecuencia fue el análisis/analítica y manejo de datos. Curiosamente, tal vez por cómo se estructura el curso de *Big Data* en las instituciones que participaron de esta investigación, se mencionó una tecnología muy específica: *Docker*,

plataforma que permite crear probar e implementar aplicaciones de todo tipo. En cuanto a la necesidad futura de un profesional de *Big Data*, existe un consenso, por cuanto este tipo de competencia será altamente requerida por la industria.

“El generar informes útiles para la empresa y clientes”. (Estudiante 16)

“Sí, creo que es un área que se encuentra en crecimiento. No necesariamente en mi ciudad (Concepción) pero sí en ciudades más grandes”. (Estudiante 5).

“Sí, porque el volumen de datos que se está generando es gigante”. (Estudiante 1)

“Sí, pienso que solo es cosa de tiempo para que implementar Big Data sea algo más recurrente”. (Estudiante 10)

“Sí, Porque el manejo de datos es muy importante”. (Estudiante 11)

“Sí, dado la relevancia de obtener valor a partir de los datos generados por las compañías”. (Estudiante 17)

Seguidamente, las preguntas que buscan conocer las expectativas que se tienen de un curso introductorio a *Big Data*, los profesores tienen un discurso similar. Todos coinciden en la utilidad de tener a futuro profesionales capacitados para asumir los desafíos del procesamiento de grandes volúmenes de datos. Cuando se pregunta sobre el tipo de proyectos en donde se requiere de los aportes de *Big Data*, se encontró una variedad de campos. Con lo expuesto, los posibles proyectos donde estos profesionales podrían participar serían: IoT, análisis de mercado, análisis histórico de clientes, detección de fraudes, pronósticos médicos y predicciones de todo tipo. Posteriormente, cuando se profundiza en esta pregunta, surge como posibilidad la realización de análisis de tipo político partidista.

“Proyectos completos con ingesta de datos, tanto batch como por eventos y explotación. Podrían proyectos más de IoT”. (Profesor 1)

“Análisis de mercado de la empresa, proyectos que tengan relación con el comportamiento histórico de los clientes, de lo que necesitan de la empresa, de los servicios que siempre demandas, de las cosas que se les podría ofrecer para crear nuevas oportunidades de negocio”. (Profesor 3)

“Análisis de preferencias de productos, análisis de fraudes, ciberseguridad, pronósticos médicos, predicciones de todo tipo, etc.”. (Profesor 4)

Finalmente, cuando se expone un escenario local, los entrevistados señalan que los posibles aportes de *Big Data* al desarrollo del país se encuentran en línea con la transformación digital, IoT, Inteligencia Artificial, desarrollo económico, desarrollo de productos/servicios que no sean del tipo *commodity*, mayor igualdad y seguridad.

“Permitir democratizar los datos, pero a la vez generar una base de conocimiento y profesionales que utilicen herramientas Word Class para el tratamiento de este tipo de información, lo que permite que el país salga de su estructura de commodities y genere productos/servicios país que agreguen valor mediante la tecnología”. (Profesor 2)

De igual forma, cuando se consulta a los estudiantes respecto a las expectativas que tienen en torno a *Big Data* y su utilidad, se mencionan muchos y muy variados proyectos. Estos van desde la limpieza de datos a la seguridad, pasando por rendimiento de la IA (Inteligencia Artificial) y *Machine Learning*, salud, aplicaciones en la misma institución de educación superior, astronomía, reciclaje y ayuda a los animales. Como es posible apreciar, la percepción es que *Big Data* tiene aplicabilidad en casi cualquier ámbito.

Del mismo modo, cuando se consulta sobre el aporte al desarrollo local, los entrevistados manifiestan que los proyectos *Big Data* estarán centrados en las áreas de salud, estadísticas nacionales, educación superior, transporte, exportación y conteo de votos. En este último caso, al igual que el segmento de profesores, queda en evidencia la escasa comprensión del concepto de *Big Data*, ya que ese tipo de procesamiento lo resuelve la arquitectura tradicional. Al momento de pedir una mayor precisión de los posibles aportes de *Big Data*, se menciona la gestión de la pandemia (COVID-19), análisis de datos, procesos industriales y toma de decisiones en políticas públicas.

Un aspecto a resaltar es que algunos de los entrevistados responden ideas muy genéricas, como mejorar el conocimiento de los datos o aumento de ventas. Nuevamente se percibe una falta de comprensión de *Big Data*, puesto que lo propuesto es posible de resolver con el análisis tradicional o bien las herramientas que aporta el *Business Intelligence* y la Minería de Datos.

“Respuestas en tiempo real y mejora de procesos industriales”. (Estudiante 1)

“Entrega oportuna de información para la toma de decisiones en implementación de políticas públicas”. (Estudiante 23)

“Mejor transparencia de datos”. (Estudiante 9)

“Un país más seguro”. (Estudiante 4)

“Bueno, depende en que se quieran enfocar, pero me imagino que en todo ámbito podría llegar a desarrollarse. No me imagino algo que no se pueda mejorar usando Big Data”. (Estudiante 8)

“Podrían aumentar ventas, ya que se sabría a qué público llegar, hacer relaciones con otros datos para así llegar al público de mejor manera”. (Estudiante 14)

Por último, dentro de los comentarios libres que incluye la entrevista, se pueden mencionar que los estudiantes manifiestan un interés por tener una mayor práctica en el uso de las tecnologías propias de *Big Data*, incluyendo un proyecto completo y la esperanza de que a futuro habrá numerosos proyectos con esta tecnología. En todo caso, la mayoría de los estudiantes indicó no tener opinión o simplemente dejó en blanco la respuesta.

Como conclusión, respecto a los estudiantes, se puede mencionar que al momento de responder la encuesta no habían logrado entender el sentido y necesidad de *Big Data*. Los comentarios son más bien desde una expectativa de conocer una tecnología que les sea útil laboralmente, aunque sin saber bien cómo y cuándo se debería utilizar.

5. Discusión

Como punto de partida, es posible señalar que tanto profesores como estudiantes exhiben una divergencia entre la definición formal de *Big Data* (López-Núñez & Cerda-Neumann, 2021) y la que ellos traen.

Respecto a los primeros, sus opiniones parecen estar influidas por su experiencia profesional. En este grupo, las expectativas sobre el curso *Big Data* se vuelven relevantes por la demanda creciente de esta competencia (Al Jafari, Zakir y Dattana, 2019).

En este contexto, cabe destacar que este primer grupo tiene la intuición de utilizar a *Big Data* como una herramienta para aumentar la competitividad de las empresas, vía una mejora de su análisis. Sin embargo esta aplicación se enmarca en las tareas asociadas a la inteligencia de negocios (Joyanes, 2019). Pese a lo anterior, es posible señalar que ellos valoran la ventaja competitiva que entrega la incorporación de la Analítica *Big Data* al proceso de toma de decisiones en la empresa moderna (Joyanes, 2019).

En tanto, los estudiantes logran construir un concepto de *Big Data* asociado a la Minería de Datos y el Machine Learning. De igual forma, relacionan *Big Data* con tecnologías de infraestructura como Hadoop, Spark y Scala. En este sentido, refiriéndonos a la infraestructura, la expectativa de *Big Data* por parte de los estudiantes, se ajusta a lo propuesto por Zheng, Z. (2021) y Tekdogan y Cakmak (2021), autores que aplican el concepto desde la infraestructura que lo soporta. Sin embargo, si bien entienden que *Big Data* se relaciona con grandes volúmenes de datos, no logran diferenciar el uso de tecnologías tradicionales, no ligadas a la computación distribuida, para el tratamiento de ellos. Bajo este escenario, la idea principal en torno a *Big Data* se relaciona con un aprendizaje de competencias ligadas al análisis de datos, sin importar el volumen de los mismos, ni menos el tipo de infraestructura necesaria para llevar a cabo dicha tarea.

En síntesis, ambas miradas coinciden en la utilidad de incluir un curso introductorio de *Big Data*. Sin embargo, los profesores son capaces de asociarlo a su experiencia laboral concreta, mientras que los estudiantes lo confunden al no entender correctamente el concepto.

6. Conclusiones

Frente a los antecedentes expuestos, los desafíos del *Big Data* en el corto plazo son enormes, tanto en la infraestructura necesaria, como también en el procesamiento en un tiempo aceptable de los datos. Esto en el sentido de la creciente generación de los mismos y la necesidad de que sean analizados. El poder de dichos datos ya es una realidad asumida por parte de los actores involucrados en esta investigación.

Un primer aporte de la investigación sería reforzar el estudio del concepto correcto de *Big Data* como punto de partida de un curso introductorio a la temática. Esto evitaría que los estudiantes lo asocien equivocadamente a otras áreas de la Informática.

Lo anterior nos lleva a sostener que el conocimiento sobre el concepto *Big Data*, como opera y cuál sería su principal objetivo, debe ser conocido por los profesionales que trabajan en torno a la gestión de los datos. Sin embargo, este estudio develó una realidad distinta a lo esperado. Se descubrió poca claridad en torno al concepto, y expectativas relacionadas con el análisis de datos en vez de los desafíos del procesamiento de grandes volúmenes de ellos.

En consecuencia, se aprecia un conocimiento del concepto *Big Data*, pero basado principalmente en la alta valoración que se asigna a los datos como fuente de información. Más que conocer *Big Data*, los informantes claves valoran el trabajo de análisis sobre los datos, sin importar el volumen, velocidad o tipo de infraestructura necesaria para su procesamiento. En esta línea, asumir como un problema resuelto los desafíos que implica la infraestructura para *Big Data* (soluciones *Cloud*), permite señalar que lo relevante no sería centrar el curso en el desarrollo de una infraestructura, sino más bien en el análisis que se realiza sobre los datos. Es decir, potenciar las competencias sobre el análisis de datos, no en el cómo gestionarlos.

Expuestos los relatos de los informantes claves, se puede decir que un curso introductorio a *Big Data* se debe centrar en el desarrollo de competencias basadas en la estadística. Por consiguiente, el futuro profesional TI debe conocer y aplicar la estadística descriptiva, como también inferencial. En esta línea, un curso introductorio a *Big Data* debe fijar como objetivo la comprensión del estudiante en torno a la estadística paramétrica, como también la no paramétrica. En lo posible, sería relevante llegar a un nivel que permita al estudiante realizar desde un análisis descriptivo e inferencial (paramétrico y no paramétrico), hasta trabajar con análisis estadístico espacial.

En líneas finales, es relevante que, tanto profesores como estudiantes, reconozcan el concepto *Big Data*, principalmente basado en sus exigencias tecnológicas de computación distribuida. Si lo anterior no se cumple, se corre el riesgo de formar profesionales que no estén capacitados en trabajar en proyectos reales relacionados con esta temática.

Referencias

- Al Jafari, A., Zakir, S. y Dattana, V. (2019). The growing demand for Big-Data Analytics. *International Journal of Information and Computing Science*, 6(5), 621 - 626.
- Althusser, L. (1970). *Ideología y aparatos ideológicos del estado*. Quinto sol.
- Bennell, P. (1996). Using and abusing rates of return: A critique of the World Bank's 1995 Education Sector Review. *International journal of educational development*, 16(3), 235-248. [https://doi.org/10.1016/0738-0593\(96\)00016-8](https://doi.org/10.1016/0738-0593(96)00016-8)
- Anshari, M., Alas, Y. y Guan, L. (2016). Developing online learning resources: Big Data, social networks, and cloud computing to support pervasive knowledge. *Education and Information Technologies*, 21(6), 1663-1677. <https://doi.org/10.1007/s10639-015-9407-3>
- Bonal, X. (2002). Globalización y política educativa: un análisis crítico de la agenda del Banco Mundial para América Latina (*Globalization and Education Policy: A Critical Analysis of the World Bank's Agenda for Latin America*). *Revista mexicana de sociología*, 3-35.
- Bowles, S. y Gintis, H. (1981). *La instrucción escolar en la América Capitalista*. Siglo XXI.
- Cao, S. (2021, march). Opportunities and Challenges of Marketing in the Context of Big Data. En *2021 Workshop on Algorithm and Big Data* (pp. 79-82).
- Cukier, K. y Mayer-Schönberger, V. (2013). The Rise of Big Data: How It's Changing the Way We Think About the World. *Foreign Affairs*, 28 - 40.
- Dede, C. J. (2016). Next steps for "Big Data" in education: Utilizing data-intensive research. *Educational Technology*. Recuperado de <https://bit.ly/31ZiEwK>
- Enterprise Strategy Group (2021). *Technology Spending Intentions Survey 2021*. Recuperado de: <https://rebrand.ly/m92z2x7>
- Espinoza, O. (2017). Neoliberalismo y educación superior en Chile: una mirada crítica al rol desempeñado por el Banco Mundial y los "Chicago Boys". *Laplace em revista*, 3(3), 93-114.
- Heyneman, S. P. (2003). The history and problems in the making of education policy at the World Bank 1960-2000. *International journal of educational development*, 23(3), 315-337.
- Heyneman, S. P. (1997). Economic growth and the international trade in educational reform. *Prospects*, 27(4), 501-530. <https://doi.org/10.1007/BF02736595>
- Hilbert, M. (2016). Big Data for Development: A Review of Promises and Challenges. *Development Policy Review*, 34(1), 135-174. <https://doi.org/10.1111/dpr.12142>
- Joyanes, L. (2013). *Big Data: análisis de grandes volúmenes de datos en organizaciones*. México D F : Alfaomega.
- Joyanes, L. (2019). *Inteligencia de Negocios y Analítica de Datos: Una visión global de Business Intelligence & Analytics*. Alfaomega.
- Kempner, K. y Loureiro, A. (2002). The global politics of education: Brazil and the World Bank. *Higher Education*, 43(3), 331-354.
- Lauglo, J. (1996). Banking on education and the uses of research. A critique of: World Bank priorities and strategies for education. *International journal of educational development*, 16(3), 221-233. [https://doi.org/10.1016/0738-0593\(96\)00015-6](https://doi.org/10.1016/0738-0593(96)00015-6)
- Linstone, H. y Turoff, M. (1975). *The Delphi method, techniques and applications*. Addison Wesley publishing.
- Long, P. y Siemens, G. (2011, September 12). *Penetrating the Fog: Analytics in Learning and Education*. *EduCause Review*, 46(5), 6.
- López-Núñez, J. y Cerda-Neumann, G. (2022). Redefiniendo Big Data: una propuesta desde la Academia. En *Mediaciones Comunicativas*, Thomson Reuters-Aranzadi.
- Maldonado, A. (2000). Los organismos internacionales y la educación en México: El caso de la educación superior y el Banco Mundial. *Perfiles educativos*, 22(87), 51-75.
- NewVantage Partners (2022). *Data and AI leadership executive survey 2022*. Recuperado de: <https://rebrand.ly/hgz4alr>
- Okoli, C. y Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Management*, 15 - 29. <https://doi.org/10.1016/j.im.2003.11.002>
- Psacharopoulos, G. (1973). *Returns to Education: An International Comparison*. Elsevier Scientific Publishing.
- Rezaee, Z. y Wang, J. (2017). Relevance of Big Data to Forensic Accounting Practice and Education: Insight from China. *7th Annual International Conference on Accounting and Finance, (Af)*, 103-110.
- Sledgianowski, D., Gomaa, M. y Tan, C. (2017). Toward integration of Big Data, technology and information systems competencies into the accounting curriculum. *Journal of Accounting Education*, 38, 81-93. <https://doi.org/10.1016/j.jaccedu.2016.12.008>
- Tekdogan, T. y Cakmak, A. (2021, august). Benchmarking Apache Spark and Hadoop MapReduce on Big Data Classification. En *2021 5th International Conference on Cloud and Big Data Computing (ICCBDC)* (pp. 15-20).

- Tójar, J. y Serrano, J. (2000). Ética e investigación educativa. *Revista electrónica de investigación y evaluación Educativa*. 6(2).
- Waller, M. y Fawcett, S. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), 77-84. <https://doi.org/10.1111/jbl.12010>
- Zheng, Z. (2021, december). Application Analysis of Hadoop in Big Data Processing. En 2021 International Conference on Aviation Safety and Information Technology (pp. 405-408)