



Documentos de Trabajo | Working Paper
CID | 2015-07

Innovación en Modelos de Negocios Basados en Datos: Los 'Big' y los 'No Tanto'



www.cid.uchile.cl

Facultad de Economía y Negocios de la Universidad de Chile

Innovación en Modelos de Negocios Basados en Datos:

Los ‘Big’ y los ‘No Tanto’

Autores:

David Díaz S. PhD.

**Departamento de Administración, Facultad de Economía y Negocios,
Universidad de Chile**

Mohamed Zaki, PhD.

**Cambridge Service Alliance. Department of Engineering,
University of Cambridge**

Working Paper – Borrador Capítulo Libro

Esta versión Mayo 2015

“Los datos son el nuevo petróleo”

(Rotella, 2012)

Abstracto

En el presente capítulo discutiremos el rol protagónico de la “(Big) Data” y cuáles son los nuevos modelos de negocios que están naciendo alrededor del análisis y procesamiento masivo de datos. En particular, en este documento nos enfocaremos desde la perspectiva privada en cómo el almacenamiento y procesamiento sostenido de datos, en mayor o menor escala, puede permitir a las organizaciones innovar en la reformulación de sus estrategias y modelos de negocios, entregando mayor valor a sus clientes por medio de nuevos productos y servicios.

1. Introducción

En la actualidad nos vemos insertos en un mundo altamente cambiante y competitivo el cual es cada vez más dependiente de las Tecnologías de la Información y las Comunicaciones (TICs) para sus procesos comerciales, sociales y políticos. Para una discusión acabada de los efectos de la adopción de TICs, el lector puede referirse al Capítulo XX

En particular, comenzaremos definiendo formalmente algunos de los términos más utilizados, y entregaremos un diagnóstico del nivel de adopción y crecimiento esperado de las tecnologías de (Big) Data en el mundo y en Latinoamérica. Hablaremos además de la innovación en los modelos de negocios basados en datos que ya se aprecian en el mundo e introduciremos un marco teórico que nos permitirá analizarlos y estudiarlos con el fin de ayudar a las empresas a posicionarse respecto de los modelos que ya existen y buscar oportunidades de desarrollo de nuevos negocios con base en aquellos aspectos que no han sido cubiertos. Finalmente concluiremos el documento entregando una discusión en torno a los desafíos para Latino América con base en lo expuesto en el capítulo.

2. Lo 'Big' y lo no tanto

Si bien el almacenamiento sistemático de datos por parte de las organizaciones no es algo nuevo, la masificación de las TICs ha generado que las tasas de crecimiento en la cantidad de data almacenada crezcan a un ritmo nunca antes visto. Existen numerosas estadísticas que se refieren a los miles de millones de bytes que se están generando y se generaran en el futuro cercano, sin embargo, para un lector poco habituado al lenguaje computacional es necesario colocar estas cifras en perspectiva. Una de las analogías más conocidas fue introducida por Erich Schmidt, Director Ejecutivo de Google, quien en el año 2010 afirmó que la humanidad actualmente genera en el transcurso de dos días el equivalente a TODA la información que fue generada por la humanidad entre el siglo IV A.C. (inicios de la protoescritura) y fines del año 2003.

Si esta cantidad de datos fuera medida en bytes estaríamos hablando de 5×10^{18} bytes, o equivalentemente un 5 seguido de 18 ceros. Tamaña cifra requiere utilizar unidades de medidas apropiadas, por lo que es común expresarla en exabyte (EB). Con esto la cifra entregada por Schmidt nos habla de aproximadamente 5 EB generados cada dos días, o equivalentemente a 900 EB generados el 2010. A partir de esta base, se estima un aumento sostenido en la cantidad de datos creados, replicados y consumidos para llegar 40,000 exabytes el año 2020 (Gantz, Reinsel, & Shadows, 2012). Quizás incluso más importante que la cantidad de datos generados es la variedad de los mismos. Diversos actores de mercado, entre ellos IBM y Oracle, estiman que el 80% de los datos generados representan contenidos no estructurados o semi-estructurados, es decir, todos aquellos datos que no son susceptibles de almacenar en un formato de bases de datos tradicionales, entre los que se incluyen las interacciones sociales (en Facebook y Twitter, por nombrar algunos ejemplos), imágenes, videos en YouTube, documentos de texto, entre muchos otros. Hablaremos más acerca de los tipos de datos en la sección siguiente. Como se puede esperar, el problema que surge de la gran variedad de formatos semi-estructurados es la dificultad que tienen los sistemas actuales para procesarlos, integrarlos y convertirlos en información relevante que pueda ser útil para la toma de decisiones.

2.1. Definiciones de Big Data

Las primeras definiciones del concepto fueron fuertemente influenciadas por el gran volumen de datos generados, sin embargo, las definiciones han ido evolucionando en el tiempo. Por ejemplo, (James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, 2011) definen Big Data como un “set de datos cuyo tamaño está más allá de las habilidades típicas de los software de datos tradicionales para capturarla, almacenarla, administrarla y analizarla”. (Gartner, 2012) la define como activos de información de alto Volumen, alta Velocidad y alta Variedad que requieren formas innovadoras y eficientes de procesamiento que habiliten su mayor comprensión y una toma de decisiones superior”.

IDC 2012 (Vesset et al., 2012) la define como “la nueva generación de tecnologías y arquitecturas de TICs que están diseñadas para extraer valor económico desde grandes bases de datos en una amplia variedad de ellos a través de habilitar la captura en alta velocidad, el descubrimiento y el análisis de los mismos”. Es posible apreciar en las definiciones, como al gran Volumen y Variedad existente en los datos, se suelen agregar variadas ‘Vs’ para describirlas. Estas incluyen a la ‘Velocidad’ con los que los datos están siendo creados, procesados y analizados –piense en los miles de transacciones *intraday* que se generan en las Bolsas de Comercio en el mundo, y la ‘Veracidad’ o calidad de los mismos. Finalmente el ‘Valor’ potencial imbuido en las mismas nos muestra como la ‘Big Data’ o la ‘Data a solas’ puede ser entendida como un Activo Estratégico en sí mismo, el cual debe ser monitoreado, administrado, protegido y hasta incluso comercializado.

Tabla 1. Big Data y Analítica Tradicional. Adaptado de Davenport 2014

	Big Data	Inteligencia de Negocios Tradicional
Tipo de Datos	Mayormente No estructurados	Con formato de Tabla, filas y columnas
Volumen de los Datos	100 terabytes o más	Menos de 10 terabytes
Velocidad o Flujo de Datos	Constante, instantáneo, alta velocidad	Estáticos o poco frecuentes
Métodos de Análisis	Analítica Avanzada (data mining)	Basado en reportes, cubos OLAP, estadística tradicional
Propósito Principal	Productos basados en Datos	Soporte de Decisiones Internas

2.2. La industria de la (Big) Data en Latino América y el Mundo

De acuerdo al pronóstico de IDC 2012 (Vesset et al., 2012) las tecnologías y servicios relacionados con tecnologías de Big Data crecerán desde \$3.2 billones de dólares en el 2010 a \$16.9 billones en el 2015. Esto representa un crecimiento anual estimado de 39.4% el que siete veces más grande que el crecimiento esperado para la totalidad del mercado de ICTs.

También de acuerdo a IDC 2014 (Llandez, Anesini, Crippa, & Villate, 2014) Latino América aún se encuentra en una fase de evangelización de las tecnologías, en la que los principales proveedores de tecnología están dedicando grandes esfuerzos en promover los beneficios de su implementación. Es así como para el año 2018 se espera que el gasto en Big Data llegue a US\$6.590 mil millones, expandiéndose un 657% en cinco años. Chile, Colombia y Perú son los países donde más se invertirá en Big Data, siendo Chile el país con mayor tasa de crecimiento de adopción de las tecnologías.

Para fines de 2014 el estudio afirma que más de 20% de las empresas medianas y pequeñas de la región estarán analizando datos semi-estructurados provenientes de sitios sociales, videos y/o datos generados por sensores. Se estima que más del 65% de los usuarios Brasileños y Mexicanos de sitios sociales gastan más de una hora por semana en Facebook, y un 24% lo hace en Twitter. Desde el punto de vista empresarial, el 56% de las empresas latinoamericanas están usando sitios sociales o de comunidades virtuales para obtener retroalimentación de sus productos y servicios.

Sin embargo, aún existe mucho espacio para crecer, dado que la adopción de estas tecnologías aún no se considera prioritario para las compañías de la región. De acuerdo con la Encuesta de Prioridades de TI 2014 de TechTarget (TechTarget, 2014), un 32% de las compañías Latino Americanas aún no tiene planes para incluir esta tecnología en el negocio, mientras que un 23.9% apenas está evaluando la opción de usar y gestionar grandes datos en su organización. Otra encuesta de opinión realizada en la región por Capgemini y referenciado por (TechTarget, 2013) muestra una foto similar en la que un 66% de los encuestados opina que la recopilación de datos impulsa la estrategia de negocios de la empresa y la toma de decisiones, pero poco más de la mitad (55%) no considera a Big Data como algo estratégico.

En este sentido, el primer desafío está asociado al nivel penetración de la (Big) Data, especialmente dadas las cifras de adopción y la baja prioridad que estas tecnologías tienen para las compañías de la región. En línea con lo mencionado, el estudio de IDC menciona que sólo el 17% de las empresas está implementando o piensa implementar Big Data en el corto plazo. Además, no lo ven como una prioridad, posicionándose bajo aquella inversión de Seguridad, ERP y Movilidad.

El segundo desafío es reconocer que lo relevante desde el punto de vista del aporte al negocio no depende del gran tamaño o volumen de la base de datos, de la rapidez con la que los datos son generados, o su gran variedad, sino más bien, el cómo estos son utilizados para innovar, ya sea reformulando procesos, reduciendo costos, o generando nuevos productos, servicios y modelos de negocios. Lo “Big” en (Big) data no tiene tanto que ver con el tamaño de la base de datos que tengo, sino en cómo puedo sacarle provecho a la información contenida en los mismos. Las empresas deben preguntarse *¿Qué tipo de cosas pueden los datos permitirme hacer que no estamos haciendo actualmente?*

2.3. El Valor de Estratégico de la (Big) Data

Desde la perspectiva de la economía global, hay quienes consideran que los datos son un nuevo factor de producción, incorporándose a la lista de factores tradicionales: tierra, trabajo y capital. Sin embargo, los datos por sí mismos no generan valor, por lo que es necesario su procesamiento para que la información que pueda ser obtenida de ellos se convierta en valor para alguien. Así, la ‘(Big) Data’ sólo puede transformarse en ‘Big Business’ en la medida que su valor estratégico sea canalizado hacia acciones y destinatarios concretos.

Desde el punto de vista empresarial, el valor estratégico de los datos se realiza en la medida que su procesamiento genere información que sea utilizada en las operaciones cotidianas a todo nivel, fomente la innovación, ayude en la toma de decisiones, permita trabajar en equipo, tanto en tiempo real como de manera asincrónica, y que en definitiva impulse el aprovechamiento del conocimiento empresarial apoyando los procesos de negocios y la relación con el medio, incluidos clientes, proveedores y *stakeholders*.

En este sentido, debe destacarse que la recolección de datos por parte de las empresas, no ha sido tradicionalmente un fin, sino más bien un sub-producto que les ha permitido crear y mantener sus ventajas competitivas. Lo clave entonces no es el Volumen y la Variedad de los datos que tenga su empresa, sino en el cómo estos datos están siendo utilizados para mejorar las ofertas que son ofrecidas a sus clientes a través de innovaciones en los productos y servicios ofrecidos, y el cómo éstos, pueden ser utilizados para generar información que permita renovar o crear desde cero nuevos modelos de negocios. La 'V' de Valor en los nuevos productos, servicios o modelos de negocios ofrecidos es por lejos la más importante.

La buena noticia es que la evidencia empírica nos muestra que aquellas empresas que adoptan una filosofía de administración y toma de decisiones soportada en el análisis de datos tienen un mejor desempeño en términos de productividad y rentabilidad (McAfee & Brynjolfsson, 2012). Si bien, en la literatura académica aún no está zanjada la discusión de cuáles son los mecanismos detrás de esta generación de valor, si está claro que los datos pueden ser utilizados para generar mejoras incrementales en los procesos y servicios. De acuerdo a (Davenport, 2014) existen cuatro maneras principales en que este Valor puede ser generado: vía reducciones de costos, vía reducciones de tiempo, vía mejoras en la toma de decisiones, y vía mejoras en los productos y servicios.

La reducción de costos es el objetivo principal de numerosas compañías, pues permite una mejora evidente en la oferta de "valor por dinero" a sus clientes. Como tal, la reducción de costos puede darse en dos dimensiones. La primera hace relación directa con la estrepitosa caída de los costos de almacenamiento y procesamiento de datos la que hace que industrias que por su naturaleza tengan que lidiar con grandes bases de datos se vean directamente beneficiadas. Ejemplo de estas son las Energéticas, Financieras, Retailers, Salud, Telecomunicaciones, Media, entre otras. La segunda dimensión hace referencia a empresas que ven reducidos sus costos como resultado complementario al haber realizado mejoras en su cadena de valor. Por ejemplo, con el objetivo de mejorar su relación con sus clientes una empresa puede realizar un estudio acabado de los patrones de compra a través de internet, identificando nuevos segmentos y potenciales mejoras de productos a dichos clientes. De paso, el reenfocar sus esfuerzos de marketing con la información generada logrará además reducir los costos por campaña, haciéndolas probablemente más eficientes y rentables.

El segundo objetivo típico de la implementación de proyecto de (Big) Data es la reducción en el tiempo necesario para realizar un proceso particular. A esto también se le conoce como reducción el "círculo de retroalimentación", es decir, el tiempo que toma una empresa en reaccionar a nueva información proveniente del entorno, implementando alguna nueva medida o acción con base en dicha información. Las compañías de Retail, por ejemplo, pueden reducir enormemente el tiempo que les toma el definir e implementar una estrategia de precios, de varios días a minutos y segundos, reaccionado rápidamente a un cambio inesperado en la demanda producto de una crisis climática.

El tercer objetivo está relacionado nuevamente con la creación de nuevos productos, servicios o modelos de negocios con base en la información extraída de los datos. Una empresa podría mejorar sus contratos de servicios e intervalos de mantención de productos industriales analizando los patrones de errores o quejas recurrentes que se aprecian en los datos. La innovación en nuevos productos y servicios "personalizados" es un área importante en la que la gran cantidad de información existente en los sitios de redes sociales juega un rol preponderante dado que le permiten a las empresas un nivel de acceso a las preferencias de sus clientes, tanto en contenido, cantidad y velocidad, nunca antes visto.

Relacionado a lo anterior, el apoyo a la toma de decisiones interna es una cuarta manera en que los datos pueden convertirse en valor agregando nuevas fuentes de datos a los diferentes modelos descriptivos y predictivos. Preguntas tales como ¿qué ofertas deberíamos ofrecerle a nuestros clientes? ¿qué clientes son los más rentables? O ¿Cuál es el mejor precio al que deberíamos vender nuestro producto? son algunas de las preguntas claves que podrán ser contestadas de manera más eficiente a través de modelos predictivos y descriptivos enriquecidos, las que luego serán convertidas en nuevos productos y servicios.

Por otra parte, nuevas realidades también pueden ser creadas en la medida que la información nos permita innovar de manera disruptiva cambiando completamente el paradigma tradicional de hacer negocios. Si se piensa, modelos de negocios en el que se cobra por las “donuts” virtuales que utiliza Homero Simpson en el juego de SmartPhone, o las “vidas” de Candy Crush en el juego de Facebook son ejemplos claros de innovación disruptiva en la manera de generar flujos de ingresos generadas por compañías que basan su éxito en la analítica de miles de datos de horas de juegos y navegación de sus usuarios.

Un quinto objetivo, no mencionado por Davenport, pero altamente importante, se refiere a la adopción de tecnologías (Big) Data como una medida de supervivencia, en el sentido, de que empresas tardías en su adopción pueden ser desplazadas por aquellos adoptadores tempranos. Las ventajas de ser el primero en generar ventajas competitivas con una determinada tecnología o modelo de negocio son altamente conocidas, y en este sentido, las empresas deben decidir ser pioneras en la implementación e innovación vía nuevas tecnologías, conllevando los riesgos que esto implica, o definirse como seguidoras adoptando la tecnología de manera rezagada únicamente como una manera de no ser completamente desplazadas por las primeras. Éstas medidas estratégicas y otros modelos de negocios basados en Datos serán discutidos con mayor profundidad en la sección siguiente.

3. Modelos de Negocios Basados en Datos ‘Big’ y no tan ‘Big’

Entonces, ¿cuáles son las posibles maneras en que se podría generar un negocio en torno a la (Big) Data?. Para contestar esta pregunta en esta sección realizaremos un recorrido por diferentes modelos de negocios que están siendo creados en diferentes partes del mundo y nos abocaremos a la tarea de tratar de clasificarlos. Para ello discutiremos un Framework o Marco Teórico recientemente propuesto por (Hartmann, Zaki, & Feldmann, 2014) en el cual se analizaron y clasificaron más de 100 modelos de negocios de compañías Start-Ups las cuales basan su núcleo de negocios en alguna manera innovadora de monetizar diferentes ofertas – productos y/o servicios – que dependen primordialmente de la explotación de datos. Cabe destacar el foco del estudio, el cual está orientado a personas del ámbito de negocios las que no necesariamente tienen un conocimiento acabado de los componentes tecnológicos. Sin embargo, éstas están principalmente abocadas a generar ofertas de valor y a administrar la comercialización de las mismas.

La elección de compañías Start-Ups como caso de estudio no es trivial, dado que de alguna manera las nuevas empresas poseen la capacidad de “partir desde cero” al definir y diseñar la manera en que estas propuestas de valor serán configuradas, lo que les da una amplia libertad a la hora de innovar en el modelo de negocio ofrecido. En este sentido, el Framework propuesto presenta un ejercicio interesante no tan sólo a quienes estén pensando en crear una nueva empresa dedicada a la explotación de datos, sino también a grandes y pequeñas empresas que quieran explorar las posibilidades de la (Big) Data.

Desde la perspectiva de Latino América, por ejemplo, el desafío propuesto es el detectar nichos de negocios (Big) Data que son esbozados en el Framework, y el cómo empresas ya existentes pueden adaptar sus modelos de negocios para incorporar componentes (Big) Data.

El hecho de que el Framework o “mapa” actual de los modelos de negocios basados en datos o – *Data-Driven Business Models (DDBM)* por sus siglas en inglés – haya sido levantado con base en firmas nacientes nos presenta además un desafío importante, dado que por su naturaleza Start-Up aún no es claro cuáles de dichos negocios serán exitosos y cuál es la forma final que adoptarán en los próximos años. Algunas estadísticas indican que por ejemplo, la tasa de fallo y adaptación de empresas Start-Ups es del orden de 90%, es decir, más del 90% de las compañías nacientes debe cambiar y adaptar – *Pivotear* en la Jerga Start-Up - su modelo de negocios inicialmente propuesto o morir en el intento. A continuación discutiremos con más detalles los temas propuestos.

3.1. ¿Qué se entiende por un Modelo de Negocios?

En la literatura académica existen numerosas definiciones de los Modelos de Negocios, y los componentes clave en su conformación, pero como es habitual en la academia, pareciera no existir un consenso definitivo respecto del tema. Uno de los primeros y más frecuentemente citados artículos fue escrito por (Chesbrough & Rosenbloom, 2002) los cuales describen el concepto de modelo de negocios de acuerdo a las funciones que desempeña dentro de la organización. Según los autores, un Modelo de Negocios es la manera en que una compañía articula las propuestas de valor que la misma ofrece a sus clientes, es decir, el valor creado para los usuarios a través de ofertas que son posibles gracias a las tecnologías que la firma utiliza. Además define un *segmento de mercado* como el destinatario de las ofertas de valor propuestas por la empresa. Así un *segmento de mercado* es definido por los autores como los *usuarios* a quienes la tecnología ofrecida es útil, cuál es su *propósito* y especifica la manera en que ésta genera *ingresos* para la compañía.

Hedman & Kalling (Hedman & Kalling, 2003) complementan la definición anterior incorporando conceptos relacionados con el ambiente competitivo en el que está inserta la empresa y además separa y define de manera más concreta los componentes y conceptos clave de un modelo de negocios. El artículo identifica siete componentes específicos: *i) los Clientes, ii) los Competidores (industria), iii) la oferta o propuesta de valor (estrategia genérica), iv) las actividades y la organización de las mismas (cadena de valor), v) los recursos y vi) la fuente de los mismos (factores de producción), y finalmente vii) los procesos longitudinales por los cuales el modelo evoluciona y se modifica.* Johnson et al (2008) Incorpora además los componentes que le permiten a la firma generar valor para sí misma y sus accionistas. La rentabilidad de la misma se genera con base en el modelo de ingresos propuesto, su estructura de costos, los márgenes obtenidos, la y eficiencia de los activos utilizados.

3.1.1. Business Model Canvas

Basado en los trabajos anteriores, quizás el trabajo más conocido y utilizado por los *practitioners* es el famoso “Business Model Canvas” o “Lienzo de Modelos de Negocios” presentado por (Osterwalder & Pigneur, 2010). En este trabajo se sintetizan los componentes o bloques de un modelo de negocios los cuales incluyen principalmente la propuesta de valor, los procesos, recursos y alianzas claves, la manera en que las relaciones con los clientes son establecidas, los canales a través de los cuales los clientes son contactados, y la manera en que todo esto genera ingresos con su estructura de costos asociada.

El Canvas está organizado de manera espacial, tal que la propuesta de valor (Value Proposition) está en el centro del mismo y constituye el “corazón” del Modelo de Negocios. A su derecha se presentan los componentes claves *externos* que tienen relación con quienes son los clientes y grupos de clientes objetivos (segmentos) de dicha propuesta de valor, y cuáles son los canales y las relaciones que son necesarias para mantenerlos. A la izquierda de la propuesta de valor se presentan los componentes claves *internos* con los cuales la empresa puede configurar y entregar la propuesta de valor, y por lo tanto representan a las actividades, recursos y alianzas claves que se requieren para conformarla. De manera *longitudinal*, pero también respetando la configuración *interna* y *externa* se presentan los componentes que soportan y reflejan el cómo se crea valor para la empresa y sus accionistas. A la izquierda abajo e internamente, se aprecian la estructura de costos que se requiere (inversiones, costos fijos, variables, etc.) y a la derecha la manera en que se monetizan dichas estructuras (Revenue Streams o Fuentes de Ingresos), por ejemplo, cobrando a través de un modelo de suscripción, licencia, pago por uso, o incluso “gratis” (el negocio se financia vía publicidad pero es libre de pago para el usuario).

3.2. Modelos de Negocios Basados en (Big) Data

Como se discutió en la sección anterior existen varias maneras en que la (Big) data puede generar valor para las compañías. Las principales incluyen la mejora y optimización de los procesos de negocios, la mejora de la relación con los clientes, la detección de nuevos segmentos de mercado, el perfeccionamiento de las comunicaciones y relaciones con proveedores, el monitoreo renovado de la competencia, la generación de nuevos flujos de ingresos, entre muchos otros.

De manera de organizar las diferentes configuraciones en que la (Big) Data puede generar valor (Otto & Aier, 2013) son uno de los primeros autores en describir modelos de negocios basados en la explotación de Datos. Ellos identifican grupos de compañías cuyo principal fin es la recolección, agregación y entrega de datos organizados de manera sistemática para que puedan ser posteriormente utilizados por otras empresas. Los autores enumeran tres tipos de empresas de acuerdo al tipo de datos que sistematizan: a) “datos empresariales” (por ejemplo, direcciones y nombres de empresas, códigos de clasificación de industrias, sectores industriales, etc.); “Datos de Contactos” (por ejemplo, números telefónicos, direcciones de e-mail de empresas y ejecutivos clave, etc.); y “Información Bancaria” (por ejemplo, número y tipo de productos bancarios contratados, situación financiera, garantías tomadas, etc.).

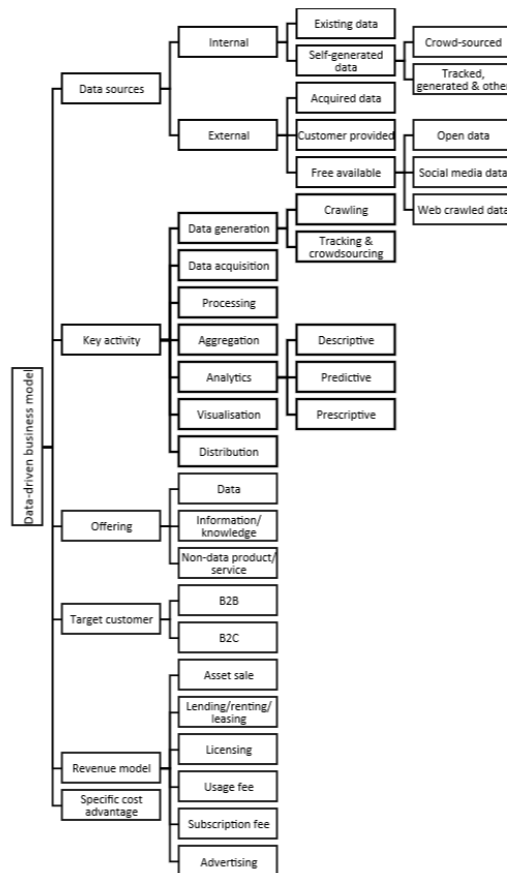
Por otra parte, un grupo numeroso de autores clasifican los modelos de negocios ofrecidos. Con base en el paradigma de la *Lógica Dominante de Servicios* los autores proponen dos tipos de modelos de negocios: “Datos-como-un-Servicio” y “Analítica-como-un-Servicio”. Los primeros incorporan aquellas empresas que agregan y proveen acceso sistemático a un variado grupo de datos públicos y privados a través de alianzas estratégicas con grandes proveedores de datos, tales como grupos Mediáticos, Bolsas de Comercio, o Grandes Proveedores de Salud como ejemplos; y los segundos incorporan aquellas empresas que proveen servicios como reportes de Inteligencia de Negocios, Análisis Automático de Texto (text mining), y Analítica Avanzada incluyendo las tareas típicas de la Minería de Datos (data Mining) tales como el modelamiento, descriptivo, predictivo y prescriptivo, en una forma tal que sus clientes puedan tener acceso fácil y directo a la información extraída de los datos para tomar decisiones de negocios.

Con esto es posible afirmar que los Modelos de Negocios Basados en (Big) Data no están circunscritos a aquellas empresas que ofrezcan servicios de analítica de datos, sino también a aquellas que se dedican a sistematizar, agregar y recolectar datos para su posterior uso por otras empresas. Un tercer tipo de

Modelo de Negocios, incluye a todas empresas que no ofrecen propuestas de valor de Datos en sí mismos, sino que más bien ofrecen nuevos productos y servicios “tradicionales” a los que se les agrega una propuesta de valor extendida con base en la información que es posible extraer de los datos. Este grupo de empresas incluye por ejemplo, empresas que ofrezcan servicios complementarios a sus productos industriales, las que podrían alertar de la necesidad de mantenimiento preventivo una turbina o motor dado el análisis de los datos que son capturados por medio de sensores durante el uso continuo de la maquinaria. General Electric, Boeing, ABB, AirBus, Caterpillar, por nombrar algunos, son ejemplos de gigantes de la industria que están transformando sus negocios tradicionales gracias al poder imbuido en el análisis de los datos.

3.3. Framework y Modelo Teórico de Modelos de Negocios Basados en (Big) Data

Los autores del Framework (Hartmann et al., 2014) proponen un completo Marco Teórico para analizar y clasificar modelos de negocios de compañías Start-Ups las cuales basan su núcleo de negocios en alguna manera innovadora de monetizar diferentes ofertas – productos y/o servicios – que dependen primordialmente de la explotación de datos. El Framework considera diferentes características y componentes modulares que en su conjunto permiten configurar una aplicación específica de Modelo de Negocios. Para ello utilizan como inspiración los componentes del Business Model Canvas de Osterwalder (2013) y como tal está compuesto por seis dimensiones principales: i) la propuesta de valor, ii) los procesos, iii) recursos y alianzas claves, iv) la manera en que las relaciones con los clientes son establecidas, v) los canales a través de los cuales los clientes son contactados, y vi) la manera en que todo esto genera ingresos dada una estructura de costos asociada.



La Figura 1 presenta el Framework y sus componentes principales. A continuación describimos con detalle cada una de ellas:

- I. **Recursos Clave:** Por definición un modelo de negocios basado en datos considera a éstos mismos su principal recurso clave. Esto no significa que sea el único recurso, pero sí que la mayoría de los otros recursos necesarios serán un soporte a la obtención, manejo y/o análisis de los mismos. Desde un punto de vista del tipo de datos que pueden ser considerados la literatura los clasifica como vimos en secciones anteriores en estructurados y no estructurados. Además, estos pueden ser clasificados según si su fuente es Interna o Externa. Las fuentes internas incluyen fuentes que ya existen o que están siendo desarrolladas en los sistemas de TICs de la empresa (sistemas CRM, ERP, SCM, etc.), y datos generados para propósitos específicos de monitoreo, ya sea de actividades en la Red, tales como el web tracking, o adquiridos por medio de sensores físicos o a través de la externalización de tareas a muchedumbres por medio de la Internet (*Crowdsourcing*). Los datos externos consideran datos que son comprados a terceros y/o provistos por socios estratégicos los cuales son por naturaleza privados, además de datos que se encuentran libremente disponibles en la Internet. Estos últimos pueden a la vez ser sub-categorizados en Open Data (datos libremente disponibles en formato electrónico y fácilmente utilizable en bases de datos), Social Media Data (provenientes de sitios sociales como Facebook), y WebCrawled data (datos que pueden ser recolectados por medio de software especializados en filtrar, seleccionar y descargar de manera eficiente los datos disponibles en Internet). Gartner (Gartner, 2012) los clasifica además en “Datos Operacionales”, “Datos Oscuros”, “Datos Comerciales”, “Datos Sociales” y “Datos Públicos”. Los datos operacionales corresponden a aquellos que provienen de sistemas transaccionales o del monitoreo de flujos de datos provenientes de sensores; los datos oscuros es aquella data que ha estado siempre presente en las organizaciones pero que actualmente no tiene un uso definido más allá de ser un registro histórico de algún evento pasado. Por ejemplo, tipos de datos oscuros son e-mails, contratos, reportes, facturas, ordenes de despacho, ordenes de trabajo, etc.; Los datos comerciales son aquellos datos que pueden ser adquiridos de una tercera parte, tales como registros de comportamiento histórico de pago, caracterizaciones socio-económicas, noticias, datos financieros, etc.; Los datos sociales son aquellos provenientes de comunidades virtuales tales como Facebook, Twitter u otros sitios similares; Finalmente los datos públicos son los provistos por organizaciones gubernamentales u otras que los ponen a disposición de la comunidad en formato abierto. Se incluyen datos económicos, censos de población, datos del tiempo atmosféricos, socio-demográficos, etc.
- II. **Actividades Clave:** Estas incluyen la “Recolección”, “Adquisición”, “Procesamiento”, “Agregación”, “Analítica”, “Visualización” y “Distribución” de Datos. Para recolectar y adquirir datos las empresas pueden generar sus propios datos internamente o pueden obtenerlos de una fuente externa. La generación puede ser hecha de varias maneras, ya sea manualmente por personal interno, automáticamente por medio del uso de sensores y herramientas de monitoreo, o a través de la externalización por medio de Crowdsourcing. El procesamiento y la agregación incluyen tareas tradicionales en la comunidad de bases de datos, tales como la extracción de datos de diferentes bases de datos, su transformación, limpieza, síntesis en tablas o reportes resumidos. La Analítica es probablemente una de las actividades más interesantes,

en el sentido que es donde los algoritmos de última generación de procesamiento de datos pueden entregarnos una comprensión valiosa de la información contenida en los datos. Esta actividad incluye la descripción de patrones en los datos y el descubrimiento de relaciones históricas en los mismos, además de permitirnos prepararnos proactivamente para el futuro por medio de técnicas predictivas y de optimización (análisis de optimización y de preparación de escenarios “Si-Entonces”) las que nos sugieren qué caminos seguir. Las técnicas modernas de data y text mining además ponen un gran énfasis en la fácil comprensión y visualización de los modelos y patrones encontrados en los datos para la toma de decisiones, por lo que visualización y distribución de resultados es otra actividad clave que puede ser entregada como propuesta de valor.

- III. **Propuestas de Valor:** Siguiendo una de las definiciones más usadas por las comunidades de Data Mining, el valor propuesto en los modelos de negocios basados en datos puede ser categorizado en tres tipos: a) datos, los que constituyen una serie de hechos o mediciones de la realidad sin una interpretación añadida, b) información y conocimiento que nace del procesamiento de los datos dándoles contexto y sentido y c) ofertas de valor complementario a productos y servicios industriales y comerciales, con base en los datos que son recolectados y analizados. Ejemplo de estos pueden ser empresas que sugieren un producto o servicio determinado, de acuerdo a los patrones de consumo y/o geolocalización de un cliente, o como se mencionó antes, empresas que ofrecen mantenimiento preventivo de acuerdo al comportamiento de la maquinaria que es posible deducir de los datos capturados por medio de sensores automáticos incluidos en ella (piense en el caso de un Automóvil o Avión Inteligente).
- IV. **Segmentos de Clientes:** De acuerdo a los autores, existen variadas formas de segmentar a los clientes, las que incluyen típicamente sus características psicológicas, demográficas, geográficas, canales de compra, conductuales, entre otras. Sin embargo, ellos utilizan la clasificación genérica, dividiendo a los clientes en empresas (B2B) y consumidores individuales (B2C), siendo posible dedicarse a atender cada segmento por separado o en conjunto.
- V. **Modelo de Generación de Ingresos:** De manera de hacer un negocio rentable cada empresa debe poseer al menos una fuente sostenible de ingresos. Típicamente estos son clasificados en siete categorías, las que no son excluyentes, es decir, una empresa puede tener una o más fuentes de ingresos: venta directa o tradicional de activos (tangibles o intangibles como los digitales); arriendo o préstamo temporal de activos, entregando momentáneamente los derechos exclusivos de uso sobre un activo en un periodo de tiempo definido; el licenciamiento, entregando la licencia de un activo a cambio de un pago; el cobro por uso de acuerdo al consumo medido en unidades por unidad de tiempo; la suscripción en la que se cobra una tarifa por el uso del activo; una comisión por la intermediación entre dos partes que se transfieren un producto y servicio; y la publicidad.
- VI. **Estructura de Costos:** Al ofrecer una propuesta de valor una empresa incurre en costos y en la utilización de sus recursos los que son limitados (tiempo, capital, trabajo, tecnología, etc.) Respecto de este punto, los autores ponen énfasis en si existe alguna ventaja relativa respecto del costo de usar, recolectar y analizar los datos. Típicamente una empresa tendrá una ventaja específica de costos si es que los datos usados en sus productos o servicios son creados independientemente (como un sub-producto) de otros bienes y servicios. Por ejemplo, una compañía que manufactura automóviles tiene dicha ventaja si captura y usa los datos que son automáticamente generados por los sistemas y sensores incluidos en la electrónica del

automóvil. En contra partida, otras empresas que deseen ofrecer servicios de datos relacionados con el tráfico, no tendrán dichas ventajas de costo relativos a los datos, pues al no ser generadores directos de los mismos deberán comprarlos a terceros, por ejemplo, a los fabricantes de autos del ejemplo anterior. Otros ejemplos clásicos de compañías con ventajas competitivas respecto del costo de producir datos son compañías de internet tales como Amazon, Google, Facebook, Twitter, entre otras, las que ya poseen inmensas cantidades de datos de variados tipos, y cuyo costo marginal de generación es tendiente a cero.

3.4. Ejemplos de Modelos de Negocios Basados en (Big) Data

Los autores reportan en su estudio las características generales de cien compañías Start-Ups que fueron analizadas en las dimensiones anteriormente mencionadas (Hartmann et al., 2014). Dentro de los resultados más importantes destaca que la mayoría de las empresas analizadas (73%) utiliza fuentes de datos externas como su recurso clave. Un porcentaje menor de ellas (16%) utiliza ambos tipos de fuentes, internas y externas, mientras que un 11% utiliza sólo fuentes internas. En términos de las Propuestas de Valor la mayoría de las compañías ofrecen servicios de analítica de datos (76%), sin embargo, sólo un pequeño número ofrece analítica avanzada, ya sea “prescriptiva” (22%) o “predictiva” (6%). Asimismo, la gran mayoría de las compañías analizadas utilizan un Modelo de Ingresos del tipo “suscripción” (62%) o basado en “pago por uso” (20%). Sin embargo, para un porcentaje bastante grande (62%) no fue posible obtener información clara acerca del tipo de Modelo de Ingresos utilizados, ya sea porque no lo divulgan, o porque aún no han sido capaces de llegar a una etapa de monetización sustentable de sus negocios. Respecto de los Segmentos de Clientes la mayoría se concentran en negocios del tipo B2B (80%), y un 70% lo hace de manera exclusiva hacia clientes corporativos. Un 13% escoge como Segmento Objetivo ambos B2B y B2C.

Al comparar los diferentes modelos de negocios de acuerdo al tipo de datos que utilizan, y al tipo de Actividad Clave en la que se enfocan es posible realizar un mapa de posicionamiento, en el que se identifican seis tipos básicos de Modelos de Negocios. De acuerdo a sus características estos fueron denominados como (ver Figura 2):

- Tipo A: “Recolector e Integrador de Datos Libres de Cobro” (‘Free data collector and aggregator’)
- Tipo B: “Analítica-como-un-Servicio” (‘Analytics-as-a service’)
- Tipo C: “Generador de Datos y Análisis” (‘Data generation and analysis’)
- Tipo D: “Descubridor de Conocimiento Gratuito” (‘Free data knowledge discovery’)
- Tipo E: “Agregador-de-Datos-como-un-Servicio” (‘Data-aggregation- as-a-service’)
- Tipo F: “Collage de datos Multi-Fuente y análisis”(‘Multi-source data mash-up and analysis’)

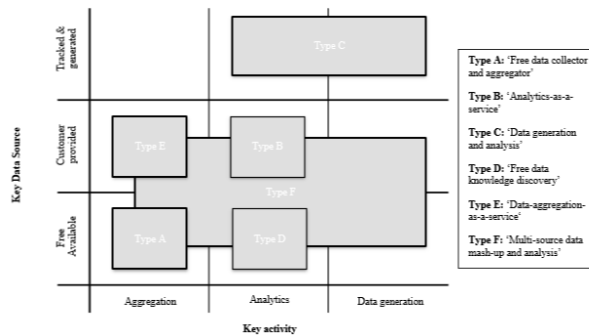


Figura 2 Tipos Genéricos de Modelos de Negocios Basados en Datos

Además fue posible identificar siete grupos o tipos de empresas los que se resumen en la Tabla 1. A continuación se presenta una descripción general de estas compañías, para una descripción más acabada refiérase al estudio original:

Tabla 2 Agrupación de Compañías con Modelos de Negocios Basados en (Big) Data

Cluster		1	2	3	4	5	6	7
Data Source	Acquired data	0	0	1	0	0	0	0
	Customer-provided data	0	1	1	0	0	1	1
	Free available data	1	0	1	0	1	0	1
	Crowd sourced	0	0	0	0	0	0	0
	Tracked, generated & other	0	0	0	1	0	0	0
Key Activity	Aggregation	1	0	0	0	0	1	1
	Analytics	0	1	1	1	1	0	1
	Data acquisition	0	0	1	0	0	0	0
	Data generation	0	0	0	1	0	0	1
Number of companies		17	28	5	16	14	6	14
Cluster Types		A	B	-	C	D	E	F

3.4.1. Tipo A: Recolector e Integrador de Datos Libres de Cobro

Las empresas de este grupo crean valor mediante la recopilación y agregación de datos de un gran número fuentes de datos, la mayoría gratuitas. Su otra actividad clave distintiva es la distribución de datos, por ejemplo, a través de una API o de un Tablero de Comando en Web (Web-based Dashboard). Otras actividades clave llevadas a cabo por empresas de este tipo son el Webcrawling (35%) y visualización (24%). Mientras que las empresas de este tipo se caracterizan por el uso de los datos libremente disponibles (100%) - en su mayoría datos de medios sociales (65%) - otros orígenes de datos, como los datos pagados (12%) o los datos generados por muchedumbre (12%) también son agregados por algunas de las empresas.

Ejemplos B2C de este tipo de empresas son *AVUXI* y *CO Everywhere*, los cuales agregan datos sobre negocios locales tales como restaurantes y bares utilizando numerosas fuentes y proporcionan esta información a los consumidores. Sus principales modelos de ingresos son la publicidad (12%) o las comisiones de corretaje (6%), además de la suscripción (47%) o el pago por uso (12%). Ejemplos B2B de este tipo es *Gnip*, fundada en 2008 y los cuales son uno de los proveedores de datos de medios sociales más grandes. Esta empresa agrega datos de una amplia gama de diferentes medios de comunicación social, ofreciendo funcionalidades de filtro y normalización de datos, además de proporcionar acceso a los datos en bruto a través de una API. Además de las fuentes sociales disponibles como Facebook o

YouTube, Gnip también es un distribuidor premium de datos de Twitter. La propuesta de valor clave de Gnip es un acceso fácil y fiable a un gran número de diferentes fuentes de datos a través de una única API. Los ingresos se generan a partir de una mezcla de cuotas de suscripción y de uso, y en función del número de fuentes incluidas y el volumen de datos procesados.

3.4.2. Tipo B: Analítica-como-un-Servicio

El segundo grupo comprende a las empresas que prestan análisis de datos como un servicio. Estas empresas se caracterizan por la realización de análisis (100%) en los datos facilitados por sus clientes (100%). Otras actividades destacables incluyen la distribución de los datos (36%), principalmente a través de facilitar el acceso a los resultados del análisis a través de una API y la visualización de los mismos (36%). Además de los datos proporcionados por los clientes, algunas empresas de este grupo también incluyen otras fuentes de datos, principalmente para mejorar la analítica. *Sendify*, por ejemplo, es una compañía que proporciona en tiempo real análisis de llamadas entrantes, al cual agrega datos demográficos externos para mejorar el análisis.

El ámbito de aplicación de los diferentes servicios de análisis incluye la detección de fraude (por ejemplo *Sift Science* o *BillGuard*); la mejora de las actividades de marketing (por ejemplo *7signal*, *Retention Science* o *Mineful*); la mejora al servicio al cliente y las relaciones (por ejemplo *Preact*, *Sendify*); el aumento de las ventas (por ejemplo *Granify*); al análisis de datos genéricos (por ejemplo *Augify*). Las empresas de este tipo se dirigen principalmente a los clientes B2B. Por lo tanto, el modelo de ingresos se basa predominantemente en comisiones de suscripción o de uso.

Si bien la mayoría de las empresas proporcionan análisis descriptivos, algunas empresas también utilizan técnicas predictivas más avanzadas. Por ejemplo, la empresa de análisis de compras en línea *Granify* trata de predecir que clientes abandonarían una compra en línea antes de completarla, mostrándoles un mensaje o código de descuento para tratar de influir en la probabilidad de compra.

3.4.3. Tipo C: Generador de Datos y Análisis

Compañías en este grupo comparten la característica de generar datos propios. Así, todas las empresas del clúster comparten la 'generación de los datos' como actividad clave. La mayoría de estas empresas también realizan análisis de estos datos. Dentro del grupo, éstas se pueden subdividir en tres grupos: las empresas que generan datos a través de crowdsourcing; Empresas de análisis web; y las empresas que generan datos a través de los teléfonos inteligentes u otros sensores físicos.

Un ejemplo de una empresa de generación de datos a través de crowdsourcing es *Swarmly*, que proporciona una aplicación de teléfono inteligente donde los usuarios pueden compartir su ubicación actual y brindar detalles de sus sentimientos acerca del lugar donde se encuentran. *Swarmly* agrega estos datos para proporcionar un mapa en tiempo real de los lugares populares, como bares, restaurantes o clubes. El segundo grupo está formado por empresas como *GoSquared*, *Mixpanel* o *Spinnaker*, que básicamente proporcionan un servicio de análisis web. Por lo tanto, recogen datos a través de un código de seguimiento integrado en los sitios web de sus clientes. Los datos son recogidos y analizados, y los informes o datos en bruto se proporcionan a través de un panel de control basado en la Web u otras interfaces. El tercer grupo son las empresas que recopilan datos a través de cualquier dispositivo físico, incluyendo los sensores de teléfonos inteligentes. Un ejemplo de una empresa de este tipo es *Automatic*, que vende un dispositivo que puede ser conectado a un puerto de datos de un vehículo y envía los datos a través de Bluetooth al smartphone del conductor. Luego recoge dichos datos

y los analiza proporcionando información sobre el estilo de conducción. Ambos modelos de negocio B2C y B2B se encuentran en este grupo. Como algunas empresas de este clúster venden dispositivos físicos para la recolección de datos (por ejemplo *Kinsa* vende termómetros para el iPhone, y *Automatic* vende un pequeño registrador de datos para el coche) algunas generan ingresos por la venta de activos físicos.

3.4.4. Tipo D: Descubridor de Conocimiento Gratuito

Las empresas de este grupo se caracterizan por el uso de datos disponible gratuitamente, sobre los cuales realizan análisis. Por otra parte, ya que no necesariamente todas las fuentes de datos gratuitas y están disponibles en un formato legible por máquinas, algunas de esas empresas rastrean los datos de la Web (generación de los datos del 50%) por medio de software de WebCrawling. Un ejemplo arquetípico de este tipo de empresa es *Gild* la que se concentra en ofrecer servicios B2B. En particular, *Gild* ofrece un servicio empresas que necesiten contratar desarrolladores de proyectos TICs, ayudándoles a reclutar programadores talentosos evaluando automáticamente el código que publican los desarrolladores en sitios de código abierto como GitHub o Google Code. Sobre la base de esta evaluación, se crea un puntaje que expresa la fuerza de un desarrollador y permite que talentos ocultos puedan ser identificados.

Aunque las empresas de este grupo son homogéneas en cuanto a las fuentes de datos clave y actividades clave utilizadas, las ofertas de valor específicas varían considerablemente. Algunos ejemplos de ofertas de valor son: el seguimiento automatizado de los sitios de revisión para hoteles (*Olerly*); recomendación de ofertas de hotel basado en el seguimiento y análisis de los diferentes sitios web de reserva (*DealAngel*); identificación de factores de influencia de los medios sociales relevantes (*Traackr*, *PeerIndex*). Ambos tipos de modelos de negocios B2B y B2C se pueden encontrar en este grupo. El tipo de análisis realizado por estas empresas también varía. Si bien la mayoría de las empresas sólo realiza análisis descriptivos, algunos utilizan técnicas analíticas más avanzadas. *TrendSpottr*, por ejemplo, trata de identificar las nuevas tendencias en los flujos de datos en tiempo real como Twitter o Facebook utilizando el análisis predictivo. *Insightpool* va un paso más allá al tratar no sólo de identificar a los clientes potenciales en las plataformas de medios sociales, sino también interactuando de manera proactiva con ellos.

Existe una variedad de modelos de ingresos dentro de las empresas de este clúster. Además de la suscripción o modelos basados en el pago por uso, las empresas B2C también dependen de los ingresos procedentes de las tarifas de publicidad o de corretaje. Por ejemplo, *DealAngel* ofrece su servicio de forma gratuita para los consumidores y recibe una comisión de los sitios web de reserva si la gente realmente reserva un hotel.

3.4.5. Tipo E: Agregador-de-Datos-como-un-Servicio

Las empresas de este grupo no crean valor mediante el análisis o la creación de datos sino a través de la agregación de datos de múltiples fuentes internas de sus clientes. Después de la agregación de los datos, las compañías presentan los datos de manera sistemática a través de diversas interfaces (distribución: 83%) y / o visualización (33%). Las áreas de aplicación se centran sobre todo en la agregación de los datos de clientes de diferentes fuentes (por ejemplo *Bluenose*) o de individuos (por ejemplo, *Who@*). Otras empresas se centran en segmentos o problemas específicos. *AlwaysPrepped*, por ejemplo, ayuda a los maestros para supervisar el rendimiento de sus estudiantes mediante la agregación de datos de los programas y sitios web educativos. Similar al Tipo B ('analytics-as-a-service),

los modelos de ingresos de estas empresas son principalmente basados en suscripción y están dirigidos al segmento B2C.

3.4.6. Tipo F: Collage de datos Multi-Fuente y análisis

Contiene empresas que agregan los datos facilitados por sus clientes con otras fuentes externas, en su mayoría gratuitas, y realiza análisis sobre estos datos. La oferta de las empresas de este grupo se caracteriza por el uso de otras fuentes de datos externos para enriquecer los datos que sus clientes les facilitan. Un ejemplo típico de un modelo de negocio de este tipo es *Welovroi*, una herramienta de monitoreo y análisis de marketing digital basado en Web que permite el seguimiento de un gran número de diferentes métricas basadas en datos proporcionados por los clientes. Sin embargo, *Welovroi* también integra datos externos y permite la evaluación comparativa del éxito de las campañas de marketing.

Otros ejemplos incluyen *FarmLogs*, el cual es un software de gestión granjas en línea, que no sólo permite el almacenamiento de los propios datos de los clientes, sino también integra datos externos, como el clima y el precio de mercado de los productos cultivados. Tales modelos se dirigen principalmente al segmento B2C y en consecuencia sus modelos de ingresos son en su mayoría de basados en suscripción.

3.5. Contribución y Brechas del Framework de Modelos de Negocios Basados en Datos

El Framework propuesto por (Hartmann et al., 2014) representa un punto de partida y mapa de referencia para el análisis y la agrupación de los modelos de negocio basados en datos existentes. Sin embargo, no pretende ser una mirada estática y definitiva de la realidad, por lo que es muy posible que en los próximos años veamos nuevas formas de generar un negocio rentable en torno a la explotación de datos.

Con todo, la contribución del Framework es mostrar a los *practitioners* las diferentes opciones al diseñar un modelo de negocio para su empresa específica y las dimensiones y las diversas características que pueden ser relevantes para su conformación. El marco permite la identificación y la evaluación de posibles fuentes de datos disponibles que se pueden utilizar en un nuevo negocio. Además, ofrece ejemplos de actividades clave, así como los modelos de ingresos más frecuentes. Los diferentes tipos proporcionan una visión sistemática de las diferentes formas de crear un negocio los cuales se pueden adaptar a una empresa o industria específica. También permite a los profesionales posicionar su negocio en un entorno competitivo y, por tanto, ofrece la posibilidad de identificar los posibles espacios de oportunidad existentes en el mercado.

4. Conclusiones y Desafíos para Latino América

En este documento hemos discutido desde la perspectiva privada cómo el almacenamiento y procesamiento sostenido de datos, en mayor o menor escala, puede permitir a las organizaciones innovar en la reformulación de sus estrategias y modelos de negocios, entregando mayor valor a sus clientes por medio de nuevos productos y servicios. Discutimos en especial, el rol de la (Big) Data y cuáles son los nuevos modelos de negocios que están naciendo alrededor del análisis y procesamiento de datos.

La primera gran conclusión que se presenta en este trabajo es que es claro que aún existe mucho espacio para crecer, especialmente dadas las cifras de adopción y la baja prioridad que estas tecnologías

tienen para las compañías de la región. La segunda gran conclusión es reconocer que el valor que es relevante desde el punto de vista del aporte al negocio no depende del gran tamaño o volumen de la base de datos, de la rapidez con la que los datos son generados, o su gran variedad, sino más bien, el cómo estos son utilizados para innovar, ya sea reformulando procesos, reduciendo costos, o generando nuevos productos, servicios y modelos de negocios. Las empresas deben preguntarse *¿Qué tipo de cosas pueden los datos permitirme hacer que no estamos haciendo actualmente?*

Para contestar dicha pregunta, se vuelve relevante el rol de las relaciones con el medio, por ejemplo, es necesario preguntarse cuáles son mis alianzas claves y mis socios principales y qué datos externos producidos en todo el eco-sistema de la cadena de valor podrían ayudarme a mejorar mi posición competitiva. Cabe recordar que la gran mayoría de las compañías Start-Ups en el ámbito de la (Big) Data utiliza fuentes de datos externas, y que sólo una porción muy pequeña de ellas utilizan tanto fuentes internas como externas. Es así, como otra de las preguntas claves es el cómo lograr que exista una integración eficiente de datos provistos por medio de alianzas, o adquiridos comercialmente, y los datos que son generados internamente en los sistemas de TICs existentes.

Los desafíos para las empresas Latino Americanas comienzan entonces con asegurar acceso a aquellos datos que puedan ser relevantes para la generación de valor, independiente de si están actualmente almacenados en mis sistemas de bases de datos existentes, o si es necesario realizar acciones estratégicas que permitan adherirse a fuentes externas. *¿Es necesario realizar una nueva alianza, un proyecto de Externalización tradicional o Crowdsourcing, o integrarnos vertical u horizontalmente para generar los datos que necesitamos?*

Relacionado al punto anterior, es necesario clarificar quien es el dueño de los datos, y si es necesario obtener permisos para hacer usufructo de los mismos. En Latino América estamos recién avanzando hacia legislaciones que protejan de manera efectiva los derechos de los usuarios respecto de sus datos, por lo que ésta es un área en la que podrían existir posibles conflictos éticos y legales, por lo que las empresas deberán poner atención especial a este punto. Además, desde el punto de vista de protección de propiedad intelectual, en ciertos casos puede ser difícil distinguir quien es el dueño de los datos que se generan y se almacenan. Por ejemplo, el lector podría cuestionarse en el caso de la empresa *Automatic* quien es el dueño de la información de los patrones de conducción: ¿Es el fabricante de automóviles, quien instaló los sensores en el vehículo?, ¿es la empresa de tecnología Satelital que provee el sistema de posicionamiento Global? ¿Es el fabricante de teléfonos inteligentes quien provee la tecnología para conectarse al satélite y capturar los datos vía *Bluetooth* para luego enviarlos a sus servidores en la nube?, ¿es el conductor, sin el cual y sin su conducción no se generaría ningún dato en primera instancia? o ¿es *Automatic* dado que sin sus algoritmos de análisis la información respecto de los patrones de conducción nunca se generaría?

Otro desafío es asegurarme que realmente entiendo el valor capturado en aquellos datos, lo que está relacionado no tan sólo con tener a disposición la tecnología necesaria para hacerlo, sino también las habilidades y la gente que sabe utilizarlos. A las habilidades técnicas de manejo de bases de datos, y el análisis avanzado de los mismos se deben agregar habilidades de negocios, ya que es necesario tomar decisiones con base en la información y el conocimiento extraído. No se gana nada con detectar un nuevo segmento de mercado, por ejemplo, si luego el área comercial no define un plan de marketing adecuado para atacarlo. Como en todo negocio, otro punto relevante es estudiar si existe alguna parte interesada para quien dicha información es valiosa, cuánto podría estar dispuesto a pagar por ella, y cuál

es la manera más de acuerdo de monetizarla, es decir, qué tipo de Modelo de Ingresos, y Estructura de Costos y Márgenes son factibles. Desde el punto de vista más técnico, pero con grandes implicancias comerciales, es el cómo asegurar la validez de los datos que se están utilizando, y sobre cuyo patrones decisiones de negocio importantes están siendo diseñadas.

El área ética es otro de los grandes desafíos para las empresas, dado que el acceso a información valiosa de nuestros clientes y usuarios puede acarrear conflictos en aquellas actividades que utilizan dicha información sin el consentimiento expreso de los usuarios. A esto además se incluyen acciones de “experimentación” las cuales son posibles de efectuar con la tecnología actual. Sitios como Facebook u OKCupid han sido ya cuestionados por realizar experimentos con los sentimientos de los usuarios, probando si por ejemplo, es posible alterar o mejorar su estado de ánimo al cambiar las noticias que se muestran o al sugerir que hacían buena pareja con otra persona.

Dado que la gran mayoría de las empresas en Latino América son PYMES, en términos de tamaño, el gran desafío es determinar cómo pueden destinar tiempo y recursos a experimentar con la (Big) Data, siendo quizás las empresas que más provecho podrían sacarle a las tecnologías dado rol su en la generación de empleo y riqueza. Éstas se podrían ver altamente beneficiadas con el acceso a tecnologías a través de la Nube, de bajo costo relativo, modelo de pago por uso y alta escalabilidad, mejorando su capacidad de innovación, eficiencia y competitividad.

Las grandes empresas y en el otro extremo las Start-Ups, poseen mayor libertad para experimentar, por lo que desde la política pública el fomento a la adopción de tecnologías (Big) data debería estar focalizado en mayor medida en aquellas empresas que no adoptarían dichas tecnologías si no contaran con dicho apoyo. Para ello, es recomendable, comenzar introduciendo poco a poco en las organizaciones las tecnologías de (Big) Data, concentrándose inicialmente en un área o producto o servicio particular, e ir creciendo y construyendo sobre éste proyecto prototipo en la medida que los resultados vayan siendo validados tanto técnica como comercialmente.

El recurso más necesario para la implementación de proyectos de (Big) Data es sin lugar a dudas la mano de obra calificada. Esto, porque se requiere un perfil especial de profesional el cual debe balancear un grupo de habilidades computacionales y tecnológicas de alta complejidad con habilidades comerciales y estratégicas que le permitan hacer uso de los *insight* que se extraen de los datos, los cuales no están egresando de las Universidades a las tasas que requiere el mercado. Por el contrario, el *hardware* y *software* puede ser más fácilmente adoptado por las empresas en la medida que se cuenten con los recursos financieros para ello, pero como cualquier herramienta empresarial avanzada se requieren los profesionales que los sepan rentabilizar.

La escasez de mano de obra especializada en (Big) data no es un tema que sólo ocurra en Chile, sino que es una tendencia global la que ha llevado a acuñar un nuevo cargo y perfil, denominado “Data Scientist”, el cual es considerado por Harvard Business Review como “el trabajo más sexy del siglo xxi”. De acuerdo a estimaciones de McKinsey Global Institute existirá una escasez de entre 140 mil y 190 mil profesionales de este tipo para el año 2018 sólo en Estados Unidos, y también una escasez de 1,5 millones de administradores y analistas que sepan cómo hacer buen uso del (Big) data.

Esto ha llevado a que dichos especialistas estén obteniendo salarios altamente atractivos y que la oferta de dichos puestos de trabajo en los mercados desarrollados siga creciendo a tasas de 2 dígitos. Asimismo, varias Universidades están creando nuevos programas y grados dedicados exclusivamente a

la formación de dichos profesionales, y existe un espacio de oportunidad para las Instituciones Educativas chilenas de ponerse al día en la oferta de profesionales que requiere el mercado.

5. Referencias

- Chesbrough, H., & Rosenbloom, R. S. (2002). The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies. *Industrial and Corporate Change*, 11, 529–555. doi:10.1093/icc/11.3.529
- Davenport, T. (2014). *Big Data@work*. Harvard Business Review Press.
- Gantz, B. J., Reinsel, D., & Shadows, B. D. (2012). *THE DIGITAL UNIVERSE IN 2020 : Big Data , Bigger Digital Shadows, and Biggest Growth in the Far East Executive Summary : A Universe of Opportunities and Challenges* (Vol. 2007, pp. 1–16). Retrieved from <http://www.emc.com/collateral/analyst-reports/idc-the-digital-universe-in-2020.pdf>
- Gartner. (2012). Big Data Analytics - Free Gartner Research. Retrieved November 08, 2014, from <http://www.gartner.com/it-glossary/big-data/>
- Hartmann, P. M., Zaki, M., & Feldmann, N. (2014). *Big Data for Big Business ? A Taxonomy of Data-driven Business Models used by Start-up Firms*.
- Hedman, J., & Kalling, T. (2003). The business model concept: theoretical underpinnings and empirical illustrations. *European Journal of Information Systems*. doi:10.1057/palgrave.ejis.3000446
- James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, A. H. B. (2011). Big data: The next frontier for innovation, competition, and productivity. Retrieved from http://scholar.google.com/scholar.bib?q=info:kkCtazs1Q6wJ:scholar.google.com/&output=citation&hl=en&as_sdt=0,47&ct=citation&cd=0
- Llandez, A., Anesini, D., Crippa, L., & Villate, R. (2014). *IDC Latin America Predictions 2014. IDC Report #LA14047*.
- McAfee, A., & Brynjolfsson, E. (2012). Big Data. The management revolution. *Harvard Buiness Review*, 90, 61–68. Retrieved from http://www.buyukverienstitusu.com/s/1870/i/Big_Data_2.pdf
- Osterwalder, A., & Pigneur, Y. (2010). *Business Model Generation. self published* (p. 281). Retrieved from <http://www.consulteam.be/media/5985/businessmodelgenerationpreview.pdf>
- Otto, B., & Aier, S. (2013). Business Models in the Data Economy: A Case Study from the Business Partner Data Domain. In *WIRTSCHAFTSINFORMATIK PROCEEDINGS* (pp. 475–489). Retrieved from <http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1029&context=wi2013>
- Rotella, P. (2012). Is Data The New Oil? Retrieved November 08, 2014, from <http://www.forbes.com/sites/perryrotella/2012/04/02/is-data-the-new-oil/>
- TechTarget. (2013). *Big Data en América Latina avanza a pasos pequeños*. Retrieved from <http://searchdatacenter.techtarget.com/es/cronica/Big-Data-en-America-Latina-avanza-a-pasos-pequenos>

TechTarget. (2014). *Encuesta de Prioridades de TI en América Latina 2014*. Retrieved from <http://searchdatacenter.techtarget.com/es/cronica/2014-Un-ano-de-proyectos-internos-de-TI>

Vesset, D., Morris, H. D., Eastwood, M., Woo, B., Villars, R. L., Bozman, J. S., & Olofson, C. W. (2012). *Worldwide Big Data Technology and Services 2012 – 2015 Forecast*.



www.cid.uchile.cl

Facultad de Economía y Negocios de la Universidad de Chile