

Universidad Nacional de San Martín

Escuela de Economía y Negocios

Utilización de *Big Data* en organizaciones. Aplicación en
startups Argentinas vinculadas al sector
agropecuario.

Tesis de Licenciatura en Administración y Gestión
Empresarial

Autor: Velázquez, Santiago

Contacto: velazquez.sgo@gmail.com

Profesor consejero: Lic. Cassin, Esteban

Director: Dr. Abella, José

Buenos Aires, julio de 2018

Resumen

El presente trabajo tiene como **principal objetivo** conocer mediante estudios de casos como se está desarrollando y desempeñando el uso de *Big Data* en *startups* argentinas relacionadas al mercado agropecuario local, con el fin de tener una perspectiva más certera sobre la configuración de los modelos de negocios basados en la tecnología de datos, para encontrar parámetros similares entre ellos y que retos deben enfrentar para ser sostenibles en el largo plazo.

La **metodología** utilizada fue la de **investigación exploratoria**, debido a que la temática a desarrollar es poco estudiada o no ha sido abordado con anterioridad y se conoce muy poco el contexto a analizar (Sampieri Hernandez, 1997). Además, ya que se pretende comprender la temática con el análisis de casos en particular, se llevó a cabo un **estudio experimental** a través de las *startups* que actualmente operan en el mercado agropecuario y que implementan *Big Data* para sus operaciones mediante encuestas de múltiples opciones.

Respecto a los resultados encontrados, se **concluye** que la gestión de los grandes volúmenes de datos ha empezado hace ya unos años a cambiar la forma de hacer los negocios, los mercados y la sociedad, convirtiéndose en un activo corporativo importante. Actualmente se vislumbra una democratización de la información con nuevas empresas, que lideradas por jóvenes profesionales generan un cambio disruptivo en los agronegocios y en la administración agropecuaria.

Los potenciales beneficios del uso de *Big Data* en el sector agropecuario son obviamente considerables, por ejemplo mediante la identificación de ineficiencias en la producción, reducción de costos y de recursos gracias al acceso de información personalizada y predictiva. Pero, simultáneamente se encuentran factores altamente limitantes como la carencia de recursos humanos capacitados y soportes tecnológicos de infraestructura adecuados.

Por último, Argentina cuenta con la oportunidad de dinamizar el mercado de *Big Data* para que estas nuevas soluciones permitan lograr mayor competitividad a nivel país. De ésta forma fomentar su estructura productiva y promover un desarrollo económico y social más sostenido y sustentable.

Palabras claves

Big Data – ventaja competitiva – *startups* – empresas – valor agregado – análisis – innovación.

Abstract

The **main objective** of the document is to know through case studies how Big Data is being developed in Argentinian startups related to the local agricultural sector. In order to have a more accurate perspective on the composition of business models based on data technology and be able to find similar parameters among them. Also to answer what challenges they must overcome in order to be sustainable in the long term.

The **methodology** used was an **exploratory research**, because the subject to be developed has not been much studied or has not been presented before and the context to be analyzed is hardly known (Sampieri Hernandez, 1997). In addition, since it is expected to understand the topic by real cases' analysis, it was executed an **experimental study** about startups that currently operate in the agricultural sector and implement Big Data for their operations through surveys of multiple choices.

Regarding the results found, it is **concluded** that the management of large volumes of data has begun a few years ago to change the way of doing business, markets, and society, becoming an important corporate asset. Currently, there is a democratization of information with new companies that, led by young professionals, generate a disruptive change in agribusiness and agricultural administration.

The potential benefits of using Big Data in the agricultural sector are obviously considerable, through for example the identification of inefficiencies in production, cost and resource reduction thanks to the access of personalized and predictive information. But simultaneously, there are highly limiting factors such as the lack of trained human resources and adequate technological infrastructure supports.

Finally, Argentina has the opportunity to boost the Big Data market so that these new solutions will make it possible to achieve greater competitiveness as a country. In this way, promote its productive structure and promote a more sustained and sustainable economic and social development.

Keywords

Big Data - competitive advantage – *startups* – companies – value added – analysis - innovation

Índice

1. Introducción.....	5
2. Marco teórico.....	9
2.1 <i>Big Data</i>	9
2.1.1 Conceptualizando <i>Big Data</i>	9
2.1.2 Fuentes de datos y su utilización	11
2.1.3 Tecnología en torno al <i>Big Data</i>	15
2.1.4 Técnicas de análisis de datos (<i>data analytics</i>).....	17
2.1.5 Inconvenientes sobre su utilización e implementación	19
2.1.6 Fases de <i>Big Data</i>	21
2.2 Utilización de <i>Big Data</i> en Argentina.....	22
2.3 Sector agropecuario	24
2.3.1 Importancia del campo en Argentina.....	24
2.3.2 <i>Big Data</i> en el sector agropecuario	26
2.4 <i>Startups</i>	28
2.4.1 Definición de <i>startups</i>	28
2.4.2 Entorno actual en Argentina	29
3. <i>Big Data</i> en <i>startups</i> agropecuarias.....	32
3.1 Sector ganadero	33
3.2 Sector agropecuario	34
4. Resultados y análisis	41
5. Conclusiones	50
6. Anexos	54
6.1 Encuesta <i>startup</i> agro <i>Big Data</i> 2018	54
7. Bibliografía.....	56

1. Introducción

El **objetivo general** del trabajo es conocer mediante estudios de casos cómo se está desarrollando y desempeñando el uso de *Big Data* en *startups* argentinas relacionadas al mercado agropecuario local.

Los **objetivos particulares** son entender cuáles son los potenciales beneficios que puede traer la implementación de esta nueva tecnología al sector tanto ganadero como agropecuario, qué características en común tienen las *startups* que llevan adelante proyectos de *Big Data* aplicados al campo argentino y por último comprender cuáles son los retos que deben vencer éstas nuevas empresas para ser sostenibles en el largo plazo.

La **estrategia metodológica** utilizada fue una **investigación exploratoria**, debido a que la temática a desarrollar es poco estudiada o no ha sido abordado con anterioridad y se conoce muy poco el contexto a analizar (Sampieri Hernandez, 1997). Además, para lograr los objetivos planteados se tomaran casos de empresas reales sobre la implementación de *Big Data*. Para ello se llevó a cabo un **estudio experimental**, ya que se pretende comprender mejor el tema con el análisis de casos en particular para analizar cómo se está trabajando en Argentina el *Big Data* y en especial atención en las *startups* del sector agropecuario. Por tal razón se entregó entre las empresas seleccionadas una encuesta para completar a través de un formulario online de diez preguntas con opciones múltiples (ver anexo 1).

Éste documento se estructura en dos grandes partes, además de ésta introducción. La **primera parte** se dedica completamente a exponer las principales resultados de la revisión bibliográfica. En primer lugar se expondrán las diferentes definiciones de expertos en el tema sobre qué es el *Big Data*, su alcance y evolución y posteriormente se puntualiza conceptualmente información sobre las *startups* con base tecnológica y su desarrollo con el contexto económico argentino. Luego, se abordan aspectos técnicos imprescindibles para la comprensión holística del tema, ya sea en cuanto a tecnología y fuentes de información necesarias para la implementación se refiere. Por otra parte, se hace mención al contexto actual del sector y con qué características se te expandiendo actualmente el mercado de datos masivos.

En la **segunda parte** luego de analizar las características de la implementación de proyectos de *Big Data* es de mera importancia llevarlo el foco de estudio a la realidad de las *startups* argentinas, específicamente en nueve empresas enfocadas en el sector agropecuario para comprender desde el interior de la organización la dinámica impuesta por el *Big Data*, su utilización y enfoque y diseño estructural. Las empresas son:

- Farmin;

- CTM Collar;
- Kilimo;
- Biobot;
- Booster Agro;
- Simapro;
- Sismagro;
- AgroZone y
- AgroApp.

El sector agropecuario fue seleccionado debido a su importancia tanto dentro de la economía nacional con su alta participación en el PBI, en el empleo y el desarrollo regional. Teniendo en cuenta además de los desafíos que se presentan en las próximas décadas para lograr la seguridad alimentaria global y al mismo tiempo por la responsabilidad que conlleva la explotación sustentable y responsable de los recursos naturales existentes.

Para finalizar el trabajo se presentan las **conclusiones** más relevantes experimentadas durante el desarrollo del presente documento.

El expresidente de Telefónica Cesar Arieta, quién desempeñó su cargo en la compañía durante 17 años, se pronunció en el 2013 durante el foro organizado por la Real Academia de Ciencias Económicas y Financieras (RACEF) de la siguiente forma: **“Los datos son el petróleo del siglo XXI**. El despliegue de sensores y el incremento de la capacidad del procesamiento, son claves en la transformación de muchos sectores y en la creación de un mundo más medible y programable”. Es por esto que en los tiempos que corren en la denominada “era de la información” los datos se han convertido poco a poco en la materia prima más deseada no solo para las grandes empresas, sino también para los gobiernos, organizaciones sin fin de lucro e incluso los mismos usuarios que navegan a través de internet.

Ésta avalancha de datos que se genera en la actualidad tiene múltiples y variadas fuentes de información, desde las más complejas y estructuradas cómo puede ser a través de satélites, teléfonos móviles, sensores o sistemas contables. Pero también, gracias a los nuevos progresos tecnológicos, es posible obtener beneficios de fuentes de datos no estructurados, cómo por ejemplo las visitas en los canales de *YouTube*, los patrones de búsqueda que ingresamos en *Google*, nuestras compras online, *posts* en redes sociales, geoposicionamiento del GPS en teléfonos móviles, solo por mencionar algunas de las actividades que realizamos en cualquier plataforma digital que deja indicios de información que hoy se considera valiosa.

Hoy en día es viable en base a los datos mencionados en el párrafo anterior, influenciar la toma de decisiones a nivel gerencial, optimizar operativamente los procesos de una organización, llevar a cabo por ejemplo una gestión de riesgos

financiera óptima, colaborar en la selección y evaluación de los recursos humanos e incluso crear nuevos modelos de negocios (IBM Institute for Business Value, 2012). Todo esto es posible gracias a la generación de un nuevo concepto denominado comúnmente como **Big Data**. Tiene la capacidad de recoger, catalogar, ordenar y analizar la gran cantidad de petabytes¹ y exabytes² de información generada. Por eso decimos que hoy un ejecutivo acompaña sus decisiones empresariales no solo con los reportes financieros, contables tradicionales sino también por ejemplo con los movimientos que sus clientes hicieron en su página web o las compras que realizaron en sus locales. Por eso se afirma que todo este volumen de información hace que sea imposible que recursos humanos puedan analizar la información al igual que lo hacen las plataformas dedicadas al *Big Data*.

A continuación, se expone lo mencionado anteriormente a través del siguiente resumen (Imagen 1) presentado en la publicación *Big Data: A survey* (Min, Chen, 2014):



Imagen 1: *The phenomenon of Big Data*. Fuente: *Big Data: a Survey* (Min Chen et al., 2014)

Pero cómo afirma Wolfram Rozas (*Business Development Executive* de *Big Data* de IBM) “un dato por sí solo no vale nada. Un dato no identifica a personas, ni marca tendencias o patrones de consumo. Pero si ese dato se recopila de forma masiva y es

¹ Petabytes: Es una unidad de almacenamiento de información cuyo símbolo que equivale a 10^{15} bytes = 1 000 000 000 000 000 de byte.

² Exabytes: Es una unidad de medida de almacenamiento de datos cuyo símbolo es el EB y equivale a 10^{18} bytes

sometido a un patrón, entonces sí nos está dando mucha información”. Actualmente partes de las grandes compañías multinacionales están inciertas y utilizan la información disponible para transformar la economía, la forma en que consumimos y la forma en que intercambiamos nuestros bienes y servicios, pero a su vez resulta difícil encontrar información y referencias sobre cómo las empresas están llevando adelante sus planes sobre *Big Data*. Sin embargo, se puede vislumbrar una tendencia y un desafío en la forma en que las compañías pueden generar valor agregado a sus modelos de negocios a través de los datos, situación que se torna compleja para empresas actuales y que no están vinculadas con las TIC (Tecnología de la Información y Comunicación). Las grandes multinacionales cuentan con una mayor cantidad de herramientas para aprovechar al máximo las nuevas virtudes del mercado. Pero también es interesante destacar el rol que van a desempeñar las nuevas *startups*, ya que se caracterizan por la flexibilidad y las condiciones de incertidumbre a las que se enfrentan (Eric, 2011).

2. Marco teórico

2.1 Big Data

2.1.1 Conceptualizando *Big Data*

Actualmente aunque ya sea un término reconocido y aceptado no existe una homogeneización sobre la definición y alcance *de Big Data*. Es por esto que vamos a tomar diferentes perspectivas de distintos autores para obtener los puntos en común entre cada uno de ellos y ver la forma en que estas definiciones se complementan.

Para comenzar, el autor Sagioglu expone que es un término que hace referencia a un conjunto de datos masivos, que tiene una estructura grande, variada y compleja, dificultando consecuentemente el almacenamiento, análisis y visualización de los procesos o resultados que se requieran obtener. Además, hace hincapié que es información útil para compañías y organizaciones para que obtengan un conocimiento más profundo y enriquecedor con el propósito de obtener una ventaja sobre la competencia (Sagioglu, 2013).

Según el artículo *Big Data: an exploration of research technologies and application cases* detalla que “*Big Data* es un término que incluye diferentes tecnologías asociadas a la **administración de grandes volúmenes de datos provenientes de diferentes fuentes y que se generan con rapidez**. [...] *Big Data* no se enfoca únicamente en tener un gran volumen de datos, sino que abarca tanto volumen como variedad de datos y velocidad de acceso y procesamiento. En la actualidad se ha pasado de la transacción a la interacción, con el propósito de obtener el mejor provecho de la información.

Para la consultora IDC: *Big Data* “describe una nueva generación de tecnologías y arquitecturas, diseñadas para extraer valor económico de un volumen muy grande de amplia variedad de datos, con una alta-velocidad de captura, descubrimiento, y/o análisis de la información”.

En adición, podemos mencionar el estudio “*2012 Big Data Work Study*” realizado por *IBM Institute for Business Value* en conjunto con la Escuela de Negocios Saïd en la Universidad de Oxford, para el que han encuestado a 1.144 profesionales de IT de 95 países diferentes para comprender cuál es el significado que los encuestados entienden por *Big Data*. Como conclusión de la encuesta realizada no se detectó una característica que se destaque sobre el resto, sino que los encuestados refirieron sus respuestas a una amplitud más grande de información, distintos tipos, fuentes y forma de análisis de los datos y también a la inmediatez de los resultados analizados. Dejando por detrás (en cantidad de respuestas) a la asociación con datos de las redes sociales o solo es una alusión a una palabra de moda. Podemos visualizar las

respuestas de todos los encuestados en la imagen número 2 (las opciones se abreviaron y normalizaron para alcanzar el 100% = 1.144 respuestas):

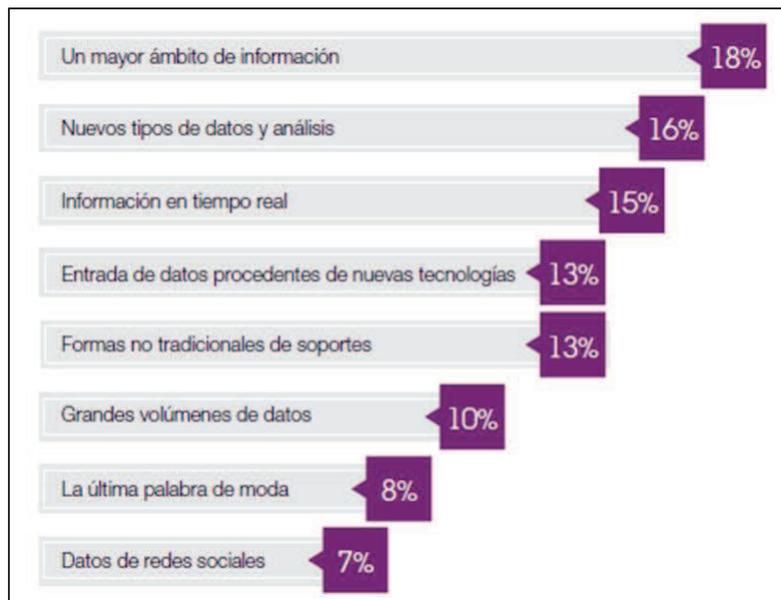


Imagen 2: Definición de *Big Data*. Fuente: IBM Institute for Business Value

En conclusión, podemos destacar que la bibliografía revisada caracteriza al *Big Data* con tres elementos principales denominado las 3 V:

- **Volumen:** hace referencia a la cantidad de datos a analizar. Debido a Ley de Moore³ no se puede definir un límite de volumen de datos necesario para considerarlo dentro de las características de *Big Data*, ya que además ésta cantidad varía según el sector en el que esta incierta la organización e incluso depende de la ubicación geográfica. Pero sí se puede definir, cómo la cantidad máxima de datos que las organizaciones utilizan para su provecho o beneficio.
- **Variedad:** distintos tipos y fuentes de datos. Tradicionalmente las empresas utilizaban registros de datos estructurados y homogéneos para la toma de decisiones, pero debido a la inserción de *Big Data* el abanico de datos y tipos de fuentes se ha ampliado a datos semi-estructurados y no estructurados y provenientes de diversas formas (audios, videos, cantidad de *clicks*, texto, *tuits*, etc.)
- **Velocidad:** consiste en el flujo de datos e inmediatez de los resultados. Hace referencia al tiempo en el que los datos son generados, capturados y

³ Ley de Moore: expresa que aproximadamente cada dos años se duplica el número de transistores en un microprocesador. En 1965, Gordon Moore afirmó que la tecnología tenía futuro, que el número de transistores por unidad de superficie en circuitos integrados se duplicaba cada año y que la tendencia continuaría. Actualmente esta ley se aplica a ordenadores personales y teléfonos móviles o celulares ya que en 1971, un chip podía contener hasta 2.300 transistores. Hoy en día caben más de 1.000 millones por unidad. (Wong Cheang, 2005)

analizados para poder utilizarlos y obtener un rendimiento incorporando datos en *streaming*⁴. Característica fundamental para procesos como la detección del fraude en el cual la variable del tiempo es pieza fundamental para mantener un sistema seguro, por ejemplo cuando *Google* sugiere palabras en tiempo real a medida que vamos escribiendo en el buscador.

En definitiva, *Big Data* es una combinación de estas características que permite principalmente capturar y analizar grandes volúmenes de información para crear nuevas oportunidades para que las organizaciones puedan generar ventajas comparativas. Estas ventajas van a variar dependiendo de la organización. Algunas de ellas ponen el foco en el almacenamiento de los datos (volumen), otras para la toma de decisiones instantánea (velocidad) y otras organizaciones los utilizan para sus modelos de negocios (variedad). Cabe destacar que ninguno de los autores revisados pone énfasis en la problemática que puede generar el exceso de datos y que tipo de problemas acarrear esta situación. Dicho escenario se desarrolla en el apartado 2.1.6.

2.1.2 Fuentes de datos y su utilización

En base a la clasificación propuesta por Barranco en referencia a las fuentes disponibles en la actualidad (es probable que tenga pronta variación con el avance de la tecnología) para ser utilizados son (Barranco, 2012):

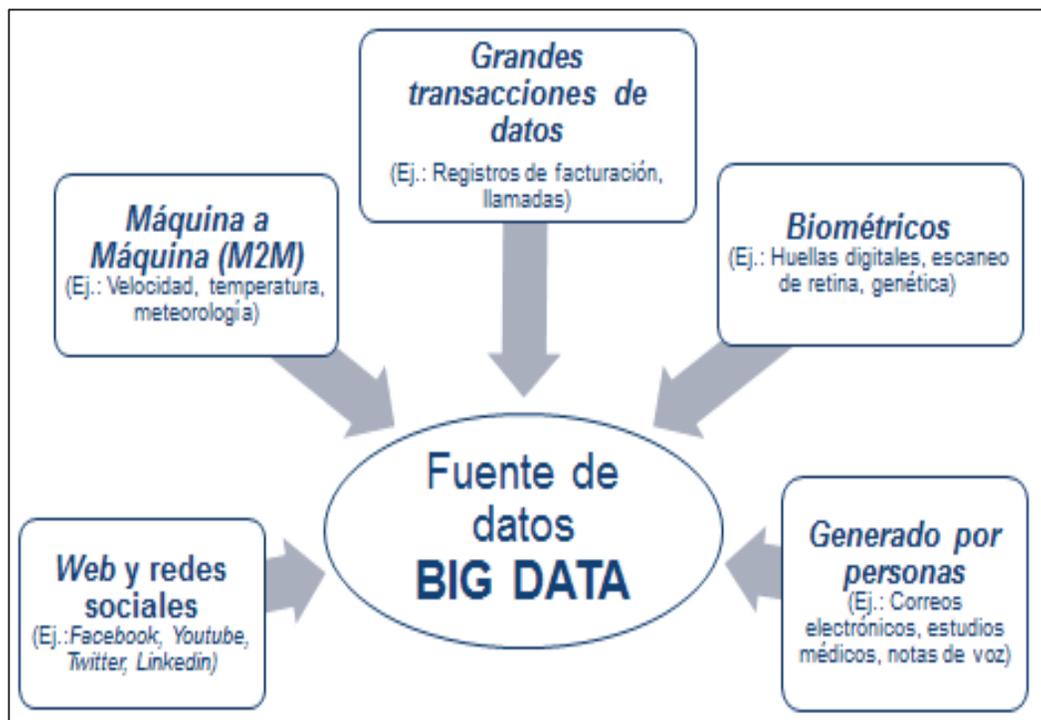


Imagen 3: Fuentes de datos *Big Data*. **Fuente:** Elaboración propia con base en el publicado por Barranco.

⁴ *Streaming*: es la distribución digital de contenido multimedia a través de una red de computadoras, de manera que el usuario utiliza el producto a la vez que se descarga.

Web y redes sociales: Para esta primera clasificación tenemos dos grandes grupos. Aquellos datos provenientes de los sitios *web* y también datos proporcionados por las redes sociales.

- Sitios *webs*: el análisis de los sitios web permite medir todos los resultados de los movimientos que se realizan en internet nutriendo de información al marketing digital dando un salto de calidad frente al marketing tradicional. Este proceso se denomina **web analytics** y consiste básicamente en el análisis cuantitativo y cualitativo del tráfico datos extraídos de un sitio web, con el fin de medir el rendimiento de la página online a través en métricas y KPI⁵ y así ofrecer una atención al cliente eficaz y mejorar la reputación online. Asimismo, es posible aprovechar estos nuevos canales para generar nuevos clientes o como herramienta de comunicación. Algunos ejemplos son:
 - Reconocimiento de marca (KPI): Mide la capacidad/conciencia que tiene un actual cliente o potencia para reconocer la marca y poder hacer alguna asociación con el producto/servicio que se ofrece. Algunas métricas pueden ser total de sesiones iniciadas, cantidad de usuarios nuevos, tráfico en el canal, etc.
 - *Engagement* (KPI): surge cuando los consumidores actuales o potenciales se encuentran atraídos por el sitio web. Algunas métricas pueden ser promedio de duración en la sesión web, porcentaje de rebote, porcentaje de páginas vistas es una sesión, etc.
 - Costo por *lead*⁶ (KPI): Hay diferentes formas de medirlo, en un sitio web puede ser por la tasa de conversión, se divide la inversión total hecha en el marketing realizado para el sitio web sobre la cantidad de *lead* obtenidos (por ejemplo cuando una persona llena un formulario, compra un producto o realiza una reserva).
- Redes sociales: Según los datos recogidos por el estudio realizado a través de la empresa *Domo* la cantidad de datos generada en internet por minuto es la siguiente:
 - *Google*: más de 2 millones de búsquedas;

⁵Métricas y KPI: Una métrica es una unidad expresada en su mayoría por unidades o porcentajes que se obtiene de herramientas de medición ya establecidas y cuyo valor prácticamente es global que, para interpretarla se requiere conocer el sistema o instrumento de medición, la escala y su significado (tasa de rebote, número de nuevas visitas al sitio o de ventas como unidades desplazadas, etcétera). Las métricas por lo general son leídas e interpretadas por el responsable de la operación ya que por medio de dicha medición se asegura la estabilidad o el cumplimiento de una acción específica.

Un KPI es la suma (o serie de métricas) que no necesariamente son universales o globales, ya que expresa claramente el desempeño de la organización por lo que su definición, lectura e interpretación son clave, los KPI se elaboran basado en métricas, pero las métricas no son KPI. Por lo que los KPI varían de negocio a negocio y de industria a industria, a diferencia de una métrica, que una vez conocida puede ser leída sin importar en qué industria o negocio se aplique. (Andes, 2015)

⁶*Lead*: es un cliente potencial que muestra interés en tu marca a través de alguna acción.

- *Facebook*: más de 684.000 artículos y 34.000 “me gusta”;
- *Twitter*: más de 236.000 *tuits*;
- Correos electrónicos: más de 100.000 *emails*;
- *iTunes*: casi 47.000 descargas de aplicaciones;
- *Tumblr*: más de 27.000 nuevos artículos;
- *YouTube*: más de 4.300 minutos de video cargados y
- *Instagram*: más de 3.600 fotos subidas.

Durante años el *marketing* se ha basado en conocer el comportamiento de los consumidores en características como la demográfica sexo, edad, ingresos, etc. Pero en la actualidad, gracias a la penetración de los dispositivos móviles y el acceso a internet, permitió al *marketing* ampliar el conocimiento y entender cómo se relacionan sus clientes actuales o potenciales. El consumo de las redes sociales por parte de los usuarios es información valiosa que permite conocer (Mejia, 2013):

- **Conocer el comportamiento del consumidor:** el marketing puede aprovechar la gran cantidad de datos de las redes sociales para conocer con precisión la conducta y preferencias de los consumidores en Internet.
- **Ajustar la comunicación con el cliente:** el análisis de la enorme cantidad de datos de las redes sociales permite ajustar la comunicación con los clientes, determinando el momento preciso y el contexto adecuado.
- **Planificar y anticipar:** el análisis de las conversaciones en Social *Big Data* permiten identificar las tendencias del consumidor, planificar nuevos productos.
- **Mejorar el servicio al cliente:** la escucha de la gran cantidad de conversaciones ayuda a prestar un mejor servicio y a integrar los canales sociales con los canales de atención al cliente, mejorando enormemente la experiencia de uso.

Máquina a máquina (M2M): El término comúnmente conocido en inglés como *machine to machine* se refiere al intercambio de una masa de datos estructurados o no estructurados entre dos dispositivos en forma remota. Para esto utiliza dispositivos como sensores o medidores que capturan algún evento en particular (velocidad, temperatura, presión, variables meteorológicas, variables químicas, sensores en geoposicionamiento, automatización, diagnóstico, etc.) los cuales transmiten a través de redes alámbricas, inalámbricas o híbridas a otros dispositivos en tiempo real para que se traduzca en información significativa, en tiempo real, con el objetivo principal de aportar valor en los servicios que se ofrecen en el negocio. Obtener las métricas generados por diferentes maquinarias/dispositivos permite capturar el comportamiento del consumidor, por ejemplo ubicación, consumo de energía o también en muchos casos se pretende obtener datos de localización, tipo, estado de objetos a través de

las tecnologías de radiofrecuencia RFID⁷ y NFC⁸ para la automatización de procesos, seguimiento en productos y control de stock, etc. No podemos dejar de mencionar en este apartado el IoT (*Internet of things*) que en esencia se habilita integrando datos de muchos dispositivos M2M diferentes y dispares, y aprovechando plataformas de software basadas en la nube para transformar esos datos en inteligencia procesable.

El IoT es un nexo de dispositivos, objetos y datos conectados a través de M2M.

Grandes transacciones de datos: son datos transaccionales originados en formato estructurados o semiestructurados pertenecientes de las transacciones que realizan las compañías a diario a través de su operatoria habitual y dependen del rubro en el que está incierta la empresa. Las compañías de *retail* recogen información relacionada sobre sus ventas (tipos de productos, gama comprada, perfil del comprador, etc.), logística (*stocks*, localización de productos, etc.) y volumen de transacciones entre otros.

Biométricos: La identificación biométrica consiste en el reconocimiento de la identidad de un usuario de una determinada red, dispositivo o aplicación tecnológica en función de rasgos físicos o de conducta que le son propios y exclusivos. El reconocimiento por voz, mediante huella digital u ocular son buenos ejemplos de identificación biométrica que han sido de gran beneficio para los sectores y agencias de seguridad e inteligencia artificial.

Generados por personas: las personas generamos diversas cantidades de datos como la información que guarda un *call center* al establecer una llamada telefónica, notas de voz, correos electrónicos, documentos electrónicos, estudios médicos, etc. Una arista interesante para comprender la capacidad que tenemos para generar datos son los dispositivos *wereables*. Son todos aquellos dispositivos que se incorporan en alguna parte de nuestro cuerpo en forma de relojes, zapatillas, prendas de vestir que logran una interacción y recopilación de datos de forma continua con la finalidad de realizar alguna función concreta como puede ser medir el estado de salud al realizar actividad física.

Pero además, cabe mencionar que existen organizaciones que gracias a las iniciativas de *open data*⁹ pueden acceder a fuentes de datos externas gracias al licenciamiento de datos. Existen empresas (*Datasift*, *Infochimps* son algunos ejemplos) que comercializan datos sobre transacciones, telecomunicaciones, logísticas y otros rubros permitiendo que diversas empresas puedan sacar provecho de esta información. Pero

⁷ RFID (*Radio Frequency Identification*): tecnología de identificación a través de etiquetas escaneadas sin necesidad de estar en contacto mediante un lector específico.

⁸ NFC (*Near Field Communication*): se trata de identificación bidireccional entre dos *chips* de forma inalámbrica para un intercambio de datos rápido.

⁹*Open data*: El concepto datos de abiertos es una filosofía y práctica que persigue que determinados tipos de datos estén disponibles de forma libre para todo el mundo, sin restricciones de derechos de autor, de patentes o de otros mecanismos de control.

también, se puede conseguir grandes volúmenes de datos de forma gratuita a través de los sitios de oficiales de instituciones gubernamentales (la Unión Europea y el gobierno de USA recopilan datos de sus instituciones nacionales) o a través de empresas privadas como *Amazon* o *Facebook* mediante *Amazon Web Services* o *Facebook Graph API*¹⁰ respectivamente.

En base a la encuesta realizada por IBM (IBM Institute for Business Value, 2012) indica que más de la mitad de los encuestados afirma que la principal fuente de datos se encuentra dentro de la organización. Esto se pueden entender ya que los datos internos son los más analizados y entendidos por los usuarios dentro de la organización.

2.1.3 Tecnología en torno al *Big Data*

Para la velocidad con que se generan los datos, la variedad de sus fuentes y los grandes volúmenes de datos anteriormente descrito en este trabajo los *softwares* y *hardwares* tradicionales quedaron obsoletos ya que no tienen la capacidad suficiente para almacenar, procesar y entregar posteriormente el servicio requerido en cada caso. Es por esto, que para el manejo de estos nuevos datos estructurados y semiestructurados fue necesario acudir a nuevas tecnologías. A continuación se presentan brevemente descripciones de los soportes tecnológicos más importantes en el mundo de *Big Data* para comprender de forma holística la temática.

Almacenamiento y procesamiento de datos: *Google* fue pionera en buscar soluciones tecnológicas que pueden solventar el gran volumen de datos con se confrontaban a diario en la compañía, ya que básicamente necesitaban un buscador de información que pudiese buscar entre todos los sitios *webs* del mundo. Es por esto, que *Google* desarrollo su propia estructura para el almacenamiento y procesamiento de datos dando origen al ***Mapreduce***.

Mapreduce es un modelo de programación para procesar y generar grandes conjuntos de datos. Los usuarios especifican una función de mapa que procesa un par clave / valor para generar un conjunto de clave / valor intermedio pares, y una función de reducción que combina todos los intermedios valores asociados con la misma clave intermedia. Muchos las tareas del mundo real son expresables en este modelo (Dean, 2004).

Finalmente se creó una versión de código abierto de *Mapreduce* llamada ***Hadoop***, es una tecnología que se encuentra en *Facebook*, *Twitter*, *eBay*, *Yahoo* y prácticamente se ha convertido en el **sinónimo de *Big Data***.

¹⁰ API: son un conjunto de comandos, funciones y protocolos informáticos que permiten a los desarrolladores crear programas específicos para ciertos sistemas operativos.

Hadoop es un sistema distribuido que corre en un conjunto de máquinas. Permite el almacenamiento masivo de datos a través de su sistema de archivos distribuido, que aprovecha la capacidad individual de almacenamiento de cada máquina de un modo conjunto. También se ha de ser capaz de procesar tal cantidad de información. Para ello, *Hadoop* implementa el paradigma de programación distribuida de *Mapreduce*, que procesa la información, explotando la capacidad de procesamiento individual de cada máquina (Prado, 2012).

Bases de datos NoSQL: Una alternativa que ha surgido para el almacenamiento de grandes volúmenes de datos y que contempla datos no estructurados o semiestructurados es el movimiento NoSQL (Not Only SQL¹¹), el cual propone el uso de sistemas que permitan gestionar estos grandes volúmenes de datos de forma eficiente y económica, cumpliendo con características como: capacidad de escalar horizontalmente y de replicar y distribuir datos a lo largo de muchos servidores; tener una interfaz simple de nivel de llamada; presentar un modelo de concurrencia débil; hacer un uso eficiente de índices distribuidos y memoria RAM para el almacenamiento de los datos y finalmente, dar la posibilidad de añadir de forma dinámica nuevos atributos a los registros de datos (Hernández Leal, 2016). Existen varias bases de datos NoSQL y varían según su forma de almacenamiento cómo Clave –Valor¹², orientadas a columnas o a documentos¹³. Algunos de ellas pueden ser *Mongo-DB*, *Redis*, *Cassandra*, *HBase* o *Amazon DynamoDB*.

Cloud Computing: Ésta tecnología busca solucionar el problema de los altos costos que implica el almacenamiento de la información y la provisión de *hardware* que para tal fin empieza a ser un inconveniente.

Servicios de *cloud computing* como *Amazon Web Services* o *Rackspace cloud*, es muy sencillo provisionar *hardware*: con un clic se pueden solicitar, configurar y arrancar cuantas máquinas sean necesarias. Una vez que no nos

¹¹ SQL (*Structured Query Language*): es un lenguaje específico del dominio que da acceso a un sistema de gestión de bases de datos relacionales que permite especificar diversos tipos de operaciones en ellos. Una de sus características es el manejo del álgebra y el cálculo relacional que permiten efectuar consultas con el fin de recuperar, de forma sencilla, información de bases de datos, así como hacer cambios en ellas.

¹² Clave –Valor: Son el modelo de base de datos NoSQL más popular, además de ser la más sencilla en cuanto a funcionalidad. En este tipo de sistema, cada elemento está identificado por una llave única, lo que permite la recuperación de la información de forma muy rápida, información que habitualmente está almacenada como un objeto binario (BLOB). Se caracterizan por ser muy eficientes tanto para las lecturas como para las escrituras.

¹³ Bases de datos documentales: Este tipo almacena la información como un documento, generalmente utilizando para ello una estructura simple como JSON o XML y donde se utiliza una clave única para cada registro. Este tipo de implementación permite, además de realizar búsquedas por clave–valor, realizar consultas más avanzadas sobre el contenido del documento. Son las bases de datos NoSQL más versátiles. Se pueden utilizar en gran cantidad de proyectos, incluyendo muchos que tradicionalmente funcionarían sobre bases de datos relacionales

sean de utilidad, podemos pararas y así prescindir del coste. Estos servicios hacen uso de la virtualización para poder ofrecer semejante flexibilidad. (Prado, 2012)

Es importante aclarar que existe un mayor número de tecnologías capaces y adecuadas para tratar un problema de *Big Data* que para otro, pero a razón de continuar con el objetivo del presente trabajo se han expuesto las principales y las que sentaron las bases iniciales de este nuevo ecosistema de trabajo contar con las herramientas principales que permitan una correcta interpretación de la temática sin ahondar en conceptos técnicos profundos que no son parte de la naturaleza del presente trabajo

2.1.4 Técnicas de análisis de datos (*data analytics*)

En este apartado se involucran para las técnicas de análisis datos campos de tipo estadístico, matemático, inteligencia artificial, entre otros para la administración y manipulación de las bases de datos recolectadas con la tecnología *Hadoop* o en cualquier otro tipo de almacenamiento. Para ello primero, nos debemos adentrar en uno de los conceptos más importantes de *Big Data* que es el **Data Mining**. La Minería de Datos combina diferentes técnicas y herramientas de los campos ya mencionados para extraer de las bases de datos de forma automática o semiautomática, conocimiento implícito, previamente desconocido, potencialmente útil y humanamente comprensible. Con el objetivo de predecir tendencias y comportamientos y/o modelos desconocidos con anterioridad (Fayyad, 1996).

Se describirán un extracto de técnicas de especial importancia para hacer accesible la comprensión de los conceptos explorados en el *Data Mining* en dos grandes campos, el *Machine learning* y el *Deep learning*. Aunque los términos se utilizan a veces como sinónimos, no son lo mismo, siendo el primero un tipo particular del segundo, es decir, el *Deep learning* es *Machine learning*, pero existen técnicas de *Machine learning* que no son *Deep learning*.

Machine learning: es una disciplina dentro de la inteligencia artificial¹⁴ que crea algoritmos que aprenden automáticamente de los datos que procesa y analiza. Cuando hablamos de “aprender” quiere decir que tienen la capacidad de identificar patrones dentro de la masa de datos y es capaz además de predecir comportamientos futuros. A mayor cantidad de datos mejor será el desempeño del algoritmo. Por ejemplo un algoritmo puede **detectar patrones de comportamiento** para predecir cuándo un cliente se va a dar de baja de una compañía de servicios de telefonía celular y gestionar la mejor acción que lo evite. Los datos históricos del conjunto de los

¹⁴ Inteligencia artificial: también llamada inteligencia computacional, es la inteligencia exhibida por máquinas. En ciencias de la computación, una máquina denominada “inteligente” percibe su entorno y lleva a cabo acciones que maximicen sus posibilidades de éxito en algún objetivo o tarea.

clientes (antigüedad, planes contratados, consumos diarios, llamadas, etc.) debidamente organizados y tratados en bloque, se puede utilizar para predecir futuros comportamientos y evitar reacciones que son perjudiciales para la compañía. Unas de las técnicas más difundidas dentro del *machine learning* es el **Clustering**. Para dicha técnica se agrupan datos de un número de clases preestablecidas o no, partiendo de criterios de distancia o similitud, de manera que las clases sean similares entre sí y distintas con las otras clases. Este método debido a su naturaleza flexible se puede combinar fácilmente con otro tipo de técnica de minería de datos, dando como resultado un sistema híbrido (Yuniet, 2009). Es ampliamente utilizado en los departamentos de *marketing* para realizar segmentaciones de mercado, para clasificar los perfiles de riesgos en clientes de seguros, en patrones de recomendaciones para clasificar gustos o preferencias, entre otros

Deep learning: lleva a cabo el proceso de *Machine learning* usando una red neuronal artificial que se compone de un número de niveles jerárquicos. En el nivel inicial de la jerarquía la red aprende algo simple y luego envía esta información al siguiente nivel. El siguiente nivel toma esta información sencilla, la combina, compone una información algo un poco más compleja, y se lo pasa al tercer nivel, y así sucesivamente. Su potencial utilidad con éxito a grandes volúmenes de datos para el descubrimiento y aplicación de conocimiento, así como a la realización de predicciones a partir de él, principalmente debido a que obtiene tasas de éxito elevadas con entrenamiento “no supervisado” (García, 2016). Los **algoritmos genéticos** forman parte fundamental del *deep learning* e intentan replicar computacionalmente comportamientos biológicos partiendo de la concepción biológica de la genética.

Se basan en aplicar los conceptos de mutaciones y combinaciones genéticas. A una población de individuos (entrada al algoritmo), se le somete a unas acciones aleatorias que los alteran de forma y manera que en la salida del algoritmo solo salen los mejores individuos en función de un criterio de selección fijado. Es uno de los tipos de inteligencia artificial más prometedores y aplicado en campos como diseño de piezas, optimización logística, previsión de comportamiento en mercados financieros, planificación de procesos, construcción de horarios y turnos (Galimany, 2014).

De las distintas herramientas mencionadas podemos extraer que todas tienen cuatro procesos en común:

- *Determinación de los objetivos:* se delimitan los objetivos en base al *Data Mining*;
- *Procesamiento de datos:* Se refiere a la selección, limpieza, reducción y enriquecimiento de la base datos;

- *Determinación del modelo:* Se comienza realizando un análisis estadístico de los datos, y después se lleva a cabo una visualización gráfica de los mismos para tener una primera aproximación. Según los objetivos planteados y la tarea que debe llevarse a cabo, pueden utilizarse algoritmos desarrollados en diferentes campos y
- *Análisis de los resultados:* Verifica si los resultados obtenidos son coherentes y los coteja con los obtenidos por los análisis estadísticos y de visualización gráfica para determinar si aportan un nuevo conocimiento que le permita influenciar sus decisiones.

El *Big Data* se sostiene bajo **dos grandes pilares**. La primera de ellas está compuesta por el **hardware** y **software** necesario para capturar, almacenar y procesar todos los datos para que estén disponibles para la generación de reportes. La segunda de ellas es la que denominamos **analítica de datos** o *data analytics*. Es la variante encargada de tomar esos datos y convertirlos en información útil. Es decir, es la encargada de otorgarle valor agregado para que las empresas obtengan su ventaja competitiva. No podemos llevar adelante un proyecto de *Big Data* sino contamos con ambos pilares, podemos tener los datos listos para utilizarme pero sin un correcto análisis se desperdician, y a la inversa, se puede contar con las herramientas de *data analytics* más avanzada pero sin una masa de datos procesada y almacenada de forma pertinente son inviables cualquier conclusión que se pretenda obtener de ese volumen de información. Para que estos dos pilares funcionen correctamente y de esta forma obtener los máximos beneficios de *Big Data* es necesario el factor clave (además de la capacidad tecnológica) que es la capacidad humana para realizar la correcta interpretación de la información que permita obtener valor de su análisis.

2.1.5 Inconvenientes sobre su utilización e implementación

Dentro de los grandes desafíos que implica esta nueva tecnología que se presenta en la actualidad involucra múltiples variables de una organización, desde la tecnología hasta los recursos humanos para transformar a los datos realmente útiles para la generación de valor agregado a los negocios de las empresas.

Según la encuesta de IBM (IBM Institute for Business Value, 2012) el entorno económico y proyectos con retornos cuantificables dejan a las compañías frenar las inversiones en tecnología de *Big Data* (razón no exclusiva de las iniciativas de esta temática). Después, según los encuestados los principales desafíos son encontrar las habilidades necesarias para que resulten operativos los proyectos, incluyendo las habilidades técnicas, analíticas y de *management*.

En particular y luego de revisar la bibliografía pertinente se desarrollan los siguientes obstáculos que toda organización debe sortear para su correcta implementación (Castro Gomez, 2018):

Desarrollo de capacidades tecnológicas: El marco tecnológico necesario para tratar con este tipo de datos es claramente más complejo que los sistemas de información habituales en cuanto a infraestructura se refiere. Teniendo en mira las restricciones presupuestarias que cada empresa tiene en particular y la dificultades de dimensionar cuantitativamente el retorno de las inversiones los costos de almacenamiento, procesamiento y de análisis son una variable importante al momento de decidir sobre la viabilidad de los programas.

Recurso humano: debe contarse con el talento capaz de definir, entender, manejar y contextualizar información de varias fuentes en varios formatos y teniendo en cuenta la reciente aparición de *Big Data*, las empresas adolecen de recursos humanos capacitados en la temática. Los perfiles indicados para explotar adecuadamente el potencial de *Big Data* deben contar con además de conocimientos técnicos en el manejo de los *software*, formación estadística y formación relacionada con el manejo de los negocios.

Identificación de datos útiles: A menudo, los datos masivos resultan confusos, de calidad variable y es aquí donde surge la implementación de la cuarta “V” desarrollada en el punto 2.1.1 sobre las características de *Big Data* y hace referencia a la **veracidad** de los datos. Dentro de la magnitud de la masa de datos analizado conseguir datos de alta calidad, es decir, fiables, exactos y pertinentes consiste en un gran desafío para llegar a obtener conclusiones acertadas y que no haya una pérdida de la inversión realizada por no llegar a cumplir con los resultados deseados.

Con respecto a la toma de muestras estadísticas y el volumen de datos (Mayer-Schonberger, 2012) explica que con ellos, muchas veces nos daremos por satisfechos con una idea de la tendencia general, en lugar de conocer un fenómeno hasta el último detalle. Además aclara que “no se renuncia a la exactitud por entero; solo abandonamos nuestra devoción por ella. Lo que perdemos en exactitud en el nivel micro, lo ganamos en percepción en el nivel macro”.

Definición de objetivos: Antes de implementar este tipo de proyectos, se debe tener en claro que enfoque se quiere aplicar y/o que procesos o áreas de la organización van a intervenir para mejorar su desarrollo. Se puede tener la información lista y depurada pero si no se conoce el fin y la utilidad que se le va a otorgar es probable que se haya perdido tiempo y grandes costos hasta que no se defina su correcta utilización.

Protección de datos: El auge de los datos masivos hace que surja nuevas problemáticas en cuanto a saber quién es el “dueño” de la información y que tratamiento le pueden dar aquellos quienes tiene accesos a esos datos. Nuevos casos surgen continuamente, brindado por un contexto en el cual el marco regulatorio es escaso e incluso nulo sobre empresas o estados que utilizan los datos brindados por

los usuarios para usos controversiales. Un caso modelo es el de *Facebook* que vendió datos de 87 millones de cuentas a la consulta *Cambridge Analytica* para su utilización en la campaña de 2017 de las elecciones presidenciales de los Estados Unidos.

Para llevar adelante proyectos de relacionados con *Big Data* es necesario como primera acción la formulación de las preguntas correctas y el planteo de objetivos claros y mensurables, administrar correctamente el actualmente escaso recurso humano y aplicar la tecnología adecuada para cada contexto. Todo esto implica un **cambio cultural** en las organizaciones que ya tienen en su concepción metodologías de trabajo ya instaladas, pero esto que en cambios abruptos se debe realizar esfuerzos mayores para integrar a las personas y a los procesos, es decir, se debe lograr que tanto los elementos tecnológicos cómo los recursos humanos pertinentes trabajen en forma conjunta.

2.1.6 Fases de *Big Data*

Para finalizar y a modo de resumen del marco conceptual sobre *Big Data* desarrollado se elaboró un esquema conceptual (imagen 4) sobre la secuencia de análisis en donde se puede encontrar siete fases lineales con un proceso de retroalimentación constante.

La primera fase, **generación**, es un proceso propio de las diversas actividad de la sociedad, en donde se genera todo el volumen de datos de naturaleza ordenada y estructurada o sin ninguna estructura pero con gran valor para las diferentes empresas que puedan obtener provecho de ellos. La segunda es la de **captura**, en la cual a través de los objetivos planteados las organizaciones extraen, de las distintas fuentes de información, los datos requeridos para cumplimentar con los requisitos y ser capaces de resolver la problemática detectada con anterioridad. Luego se necesita una **limpieza** de estos datos recolectados (aquí es donde entra en juego el *Data Mining*), ya que presentan en su concepción redundancia o datos que son inútiles para los fines propuestos y que si no se detectan a tiempo incrementa los tiempos de análisis y almacenamiento posteriores, aumentando por consecuencia los costos de operación.

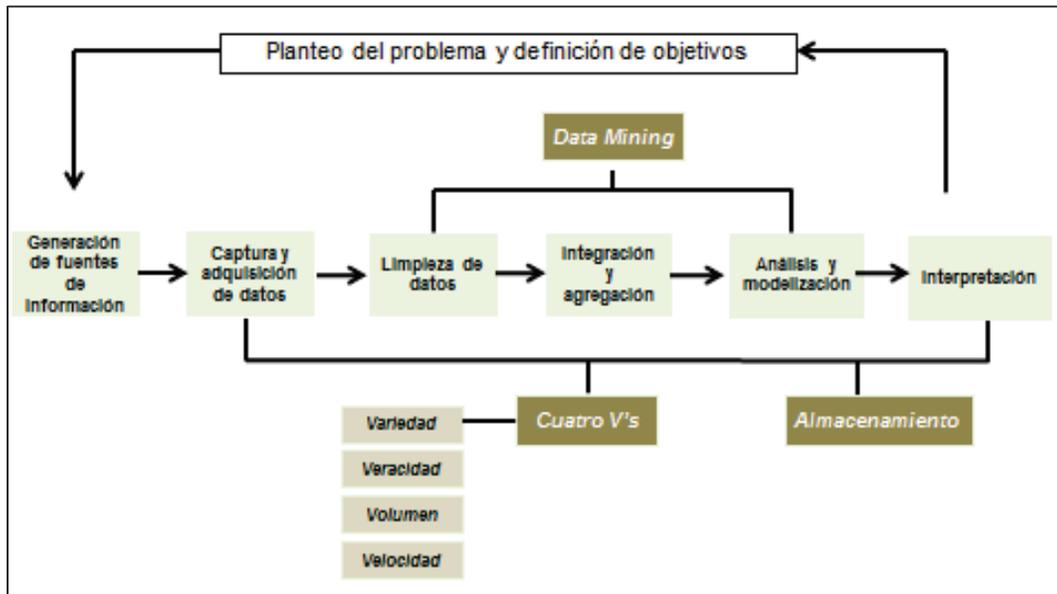


Imagen 4: Fases de Big Data. Fuente: Elaboración propia.

Después es necesaria la **integración** de los datos. Aquí se preparan los datos para ganar agilidad, rapidez y facilidad de uso para el **análisis** final en el cual se debe extraer rápidamente información desde los datos masivos para generar valor a las organizaciones y ayudar en la toma de decisiones. Para finalizar, es aquí donde se requiere la mayor capacidad de recursos humanos profesionales capacitados para la **interpretación** de los resultados.

2.2 Utilización de Big Data en Argentina

Para este apartado nos basaremos en la encuesta realizada por el CIECTI (Centro Interdisciplinario de Estudios en Ciencia, Tecnología e Innovación) denominada "Encuesta digital: Descubriendo *Big Data* en Argentina" que consistió en la realización de dos encuestas, una destinada a los grupos de investigación que se desempeñan en instituciones científicas y tecnológicas y otra dirigida a empresas (todas vinculadas a actividades de *Big Data*). Se obtuvieron respuestas completas de 15 grupos de investigación y de 18 empresas, representando el 75% y el 51% de los padrones, respectivamente.

Los principales resultados de la encuesta se pueden resumir en los siguientes tópicos:

- **Complejidad** de los proyectos de *Big Data*: el 63% (sobre 16) de las empresas respondieron que es de complejidad media; el 25% respondió que es de baja complejidad y sólo una empresa la considera de alta complejidad. Las principales razones brindadas por los participantes **remiten al volumen de datos generados, el desconocimiento sobre el tema y el desarrollo incipiente de proyectos**. Los posibles usuarios del análisis de *Big Data* desconocen de qué se tratan estos desarrollos y los potenciales beneficios,

mientras que las empresas que pueden ofrecer servicios relacionados encuentran dificultades a la hora de mostrar el retorno de las inversiones de los proyectos;

- **Destino** de productos y servicios: el 46% de los clientes son empresas privadas nacionales y otro 46% son firmas extranjeras. La alta participación de empresas privadas extranjeras en la facturación de un sector emergente indicaría la existencia de capacidades locales competitivas orientadas hacia el exterior por **falta de un mercado interno suficientemente demandante**;
- **Origen** de los proyectos: el **sesgo privado de la actividad** o bien la **escasa participación del sector público** en la demanda de proyectos de *Big Data*. Se trata de un área nueva y la producción de grandes volúmenes de datos todavía es incipiente en el sector público o se desconocen las potenciales aplicaciones de los mismos: el nivel de incertidumbre sobre *Big Data* estaría limitando la demanda de estos productos. Por otro lado, se trataría de un problema institucional: la característica transversal de este subsector afecta las actividades de diferentes organismos públicos, lo cual requiere de coordinación y comunicación entre los mismos.

En conclusión, el *Big Data* en Argentina posee una **demanda escasa** y de **baja complejidad**, explicado por la **falta de conocimiento sobre la temática, problemas institucionales** y **limitaciones en la infraestructura**. Estos resultados indican la existencia de margen para la **intervención dinamizadora del Estado**.

Sin embargo, cabe destacar que recientemente, el Ministerio de Comunicaciones de la Nación Argentina creó un nuevo ente llamado **Observatorio Nacional de Big Data** con miembros *ad honorem* con el principal objetivo de la promoción y creación de nuevas plataformas de *Big Data* dentro y fuera del Estado.

También en la Ciudad Autónoma de Buenos Aires a través del **Ministerio de Modernización** está llevando adelante proyectos como el Mapa de Oportunidades Comerciales o la Red de Sensores Ambientales en los cuales ofrecen un servicio a la comunidad a través del análisis de datos disponibles por el gobierno de la Ciudad. En el primer caso, es una plataforma para que los emprendedores puedan detectar los riesgos de abrir un comercio e identificar las oportunidades del mercado. Permite acceder a una evaluación comercial sobre un barrio, a través de un reporte detallado con el número de aperturas y cierres de locales en los últimos dos años, información del promedio de ventas de esos comercios y un índice de estabilidad que se basa en la supervivencia de los negocios a lo largo del tiempo. Además, muestra diversos indicadores poblacionales como la cantidad de personas que viven y trabajan en una zona en particular, desagregados por rango etario y sexo, y datos de inmobiliarias que indican el valor promedio del alquiler y la venta del metro cuadrado. Respecto al segundo proyecto, es una plataforma de 1000 sensores colocados por la Ciudad de

Buenos Aires que detectan cuestiones atmosféricas y de calidad ambiental, además de niveles de ruido, radiación solar y otras variables relacionadas al comportamiento urbano, como flujos de personas o autos.

2.3 Sector agropecuario

2.3.1 Importancia del campo en Argentina

En un país periférico como la Argentina en el cual sus modelos económicos (sin importar a que gobierno pertenezcan) tiene como pieza fundamental para su funcionamiento la restricción externa de divisas y es así como el sector agropecuario ha sido desde tiempos remotos el más importante para el desarrollo del país a través de **su aporte de divisas con las exportaciones** de materia primas. Además de ser un **generador de empleo** a lo largo y ancho del país de todas las cadenas agroalimentarias también es un importante **contribuyente a las arcas del Estado** a través de diferentes impuestos.

La producción asociada al campo tiene una extensa cobertura geográfica en todo el país, siendo la actividad más federal dependiendo de estos en gran medida de la mayoría de los recursos naturales de la Argentina.

En consecuencia, según el último informe del FADA (Federación Agropecuaria para el desarrollo de la Argentina) del año 2017 cuando Argentina produjo 122 millones de toneladas de granos, 5 millones de toneladas de carne (aviar, vacuna y porcina), 9.895 millones de litros de leche, y otras producciones tales como biocombustibles, vinos y cultivos no tradicionales el sector produjo (Pisani & Miazza, 2017):

- **1 de cada 6 puestos** de trabajo privados (directos e indirectos);
- **1 de cada 10 pesos** del Producto Bruto Interno;
- **1 de cada 10 pesos** de la recaudación tributaria nacional de AFIP y
- **7 de cada 10 dólares** de las exportaciones totales del país.

Entre los indicadores que reflejan el impacto e importancia del sector en Argentina se detallan en el presente informe: generación de empleo (tanto directo como indirecto); impacto de la actividad en el PBI nacional; el aporte tributario y la generación de divisas por exportaciones.

Generación de empleo: La generación de empleo es el primer, y más importante, efecto multiplicador del desarrollo de las cadenas agroalimentarias, por ejemplo, la producción de lácteos tiene un coeficiente multiplicador de empleo de 6,10; la matanza de ganado 5,52; y el curtido y terminación de cueros 4,49 y que en comparación con la producción de automotores que tiene un 4,66 y la fabricación de textiles 1,72 2 (Regunaga & Tejeda, 2015). Las cadenas agroalimentarias generaron 2,5 millones de puestos de trabajo para el 2016, **el 17% del empleo total nacional (privado)**. Si se

suma el empleo público, dicho porcentaje se reduciría al 14%. Del total de puestos generados, **el sector primario es el mayor generador de empleo, abarcando el 46% del total**, 26% lo genera el eslabón comercial, 16% la industrialización de las diferentes producciones primarias, siendo el restante 12% generado por el transporte (Pisani & Miazzo, 2017).

Producto Bruto Interno: Las cadenas agroalimentarias aportaron al PBI del año 2016 de un total de aproximadamente 8 billones de pesos el 10,4% del total del PBI, equivalente a \$ 840.035 millones. Este aporte se puede dividir en dos grandes eslabones: el sector primario (PP) y la industria manufacturera de origen agropecuario (Gráfico 1) (Pisani & Miazzo, 2017).

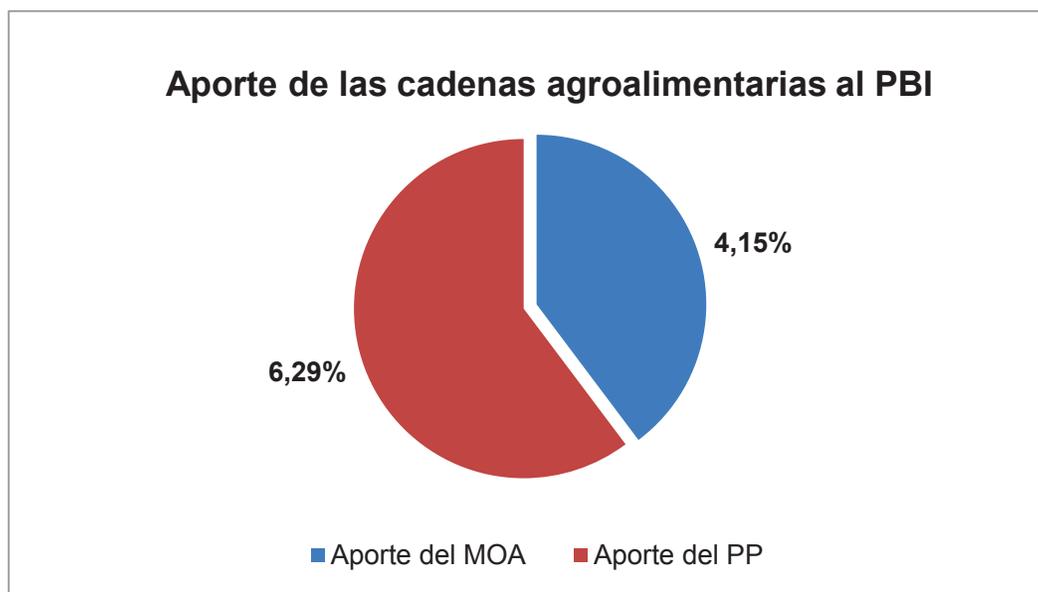


Gráfico 1: Aporte al PBI Fuente: FADA en base a INDEC

Aporte tributario: Para el año 2016, Argentina tuvo una recaudación tributaria de \$ 2.273.047 millones. Que las cadenas agroalimentarias aportaron \$ 208.883 millones, de los cuales \$ 84.445 millones fueron aportados por el sector agropecuario primario y \$124.438 millones por la industria agroalimentaria. Sobre la recaudación nacional, las cadenas agroalimentarias aportaron el 9,2%, pero asciende al 10,7% si solo se consideran la recaudación de los impuestos que fueron medidos para las cadenas agroalimentarias. Los impuestos tomados en cuenta para el cálculo son: Ganancias, Ganancia Mínima Presunta, Seguridad Social (CSS), Derechos de Exportación (DEX), IVA e impuesto a los Créditos y Débitos (ICDB) (Gráfico 2) (Pisani & Miazzo, 2017).

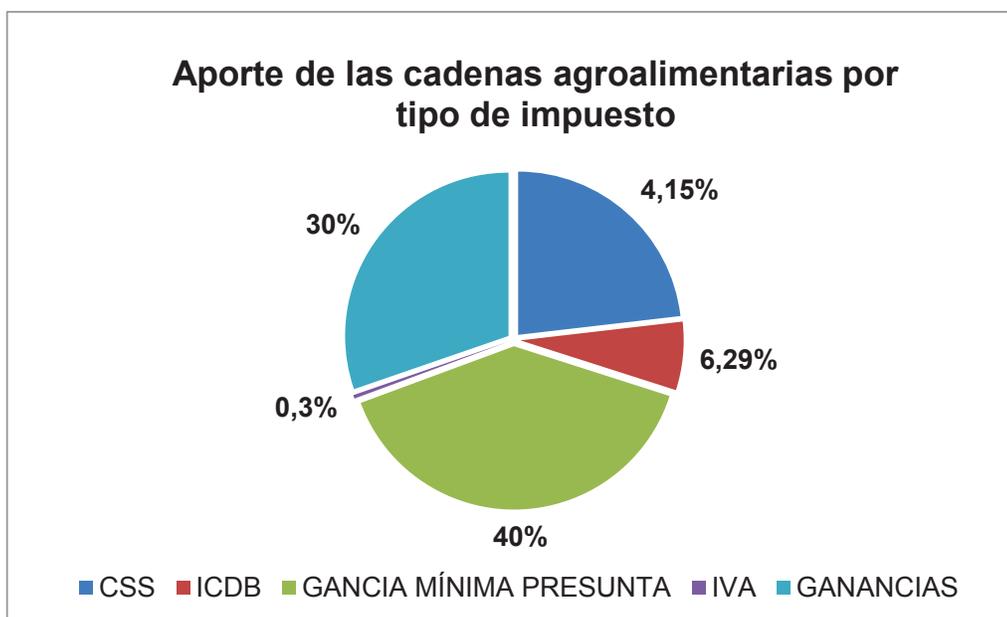


Gráfico 2: Aporte según tipo de impuesto **Fuente:** FADA en base a INDEC

Exportaciones: Para el año 2016 el campo argentino aportó el 66% del total de divisas por exportación, de los cuáles USD 14.677 millones corresponden a productos primarios, mientras que USD 23.342 millones fueron explicados por exportaciones de manufacturas de origen agropecuario. Además, para el mismo año no sólo fue el sector que mayor aporte de divisas realizó al Estado argentino, sino que presentó una balanza comercial con superávit. Teniendo en cuenta individualmente las balanzas comerciales del resto de los rubros, todas ellas, salvo las cadenas agroalimentarias, presentan déficit (Pisani & Miazzo, 2017).

2.3.2 *Big Data* en el sector agropecuario

Las actividades agropecuarias con el afán de mejorar sus niveles de eficiencia y expandir su productividad y así mejorar el aporte tanto económico como social al país, teniendo en cuenta la escasez de recursos naturales como el agua, tierras de cultivos, energía, se ven obligadas a constantemente **incorporar tecnologías y conocimiento que ayude a fortalecer los sistemas productivos** para generar mejores condiciones de competitividad internacional a través de nuevas herramientas.

Argentina ha crecido en la producción de granos en los últimos 18 años a razón de 3,7 M/toneladas/año, gran parte de este aumento productivo se dio por la generación y adopción de nuevas tecnologías (INTA, 2011). En la actualidad hay una serie de desarrollos y avances vinculados con la tecnología agropecuaria que han marcado el rumbo de los sistemas informáticos. En términos genéricos, se pueden apreciar los siguientes usos:

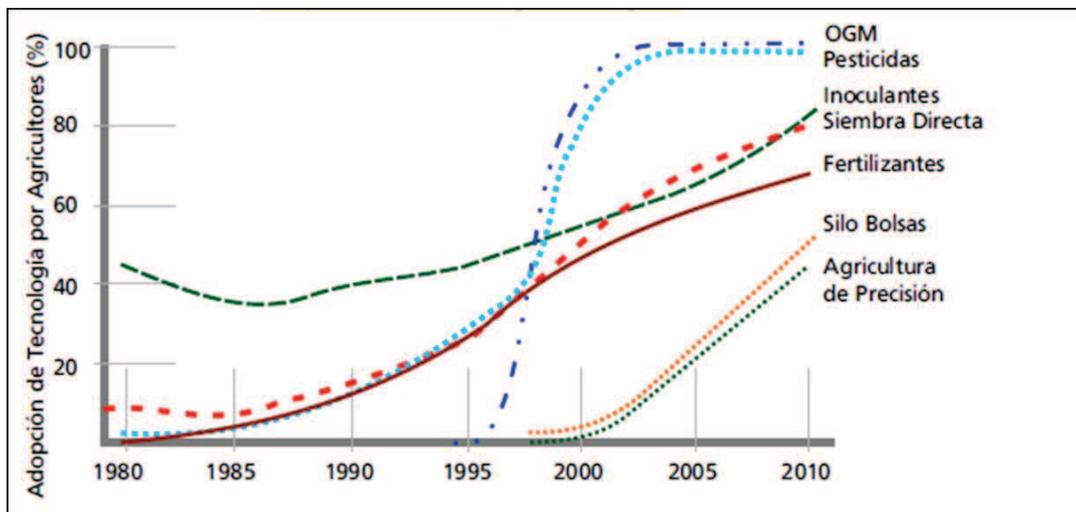


Gráfico 4: Adopción de la tecnología en Argentina. **Fuente:** INTA

Las explotaciones agropecuarias en toda su cadena de valor generan una gran cantidad de datos (proviene de estaciones meteorológicas, sensores de suelo, sistemas de riego, maquinaria e instalaciones, datos de sistemas informáticos, ERP, reportes, documentos legales, laboratorios, predicciones meteorológicas, imágenes de satélites y drones, observaciones en campo, etc.) con capacidad de ser procesados y convertidos en información útil. Pero usualmente estos datos se almacenan en sistemas desconectados, donde crecen exponencialmente en volumen y se hace más difícil usarlos con eficiencia y la frecuencia deseada. Es por esto que actualmente los usuarios (propietarios de campos, granjas, centros de transformación y procesado, almacenes, invernaderos) que dependen de estos datos deben consultar múltiples sistemas, para la correcta toma de decisiones.

Para que el productor reduzca la incertidumbre en la toma de decisiones (por ejemplo para minimizar el riesgo de las condiciones meteorológicas inciertas, plagas, enfermedades, la variación de los precios y optimizar los niveles de producción) con los datos mencionados en el párrafo anterior se están desarrollando proyectos, que realizan predicciones agrícolas y de mercado con el fin de asesorar a productores con información fiable referente a cultivos, suelos, productos y climas a través de *Big Data*. Es entonces aquí en donde se aprecia una mejora palpable en la utilización de *Big Data* en el sector agropecuario. Los ámbitos de aplicación actualmente son (Serrano, 2017):

- En decisiones de **inversión agrícola** a largo plazo o en planificación de la producción a corto / medio, analizando tendencias y evoluciones inerciales;
- En la **agricultura de precisión**, haciendo mejores previsiones de cosecha, mejorando y anticipando la probabilidad de ocurrencia de enfermedades o plagas, reduciendo impactos ambientales;

- En **transformación/manipulación de productos agroalimentarios**, haciendo una mejor segmentación de los mismos, mejorando la calidad, planificando y optimizando los recursos necesarios en el proceso de transformación;
- En **logística y distribución de productos agroalimentarios**, analizando mejor los procesos logísticos y optimizando costes y calidad de servicio;
- En **comercialización y marketing de productos agroalimentarios**, conociendo y segmentando mejor la demanda, alineándola con la oferta, anticipando y gestionando a nuestro favor las oscilaciones e imperfecciones del mercado, analizando tendencias y
- En un **sistema de precios inteligente** más cercano al tiempo real y mejor adaptado para lograr un mejor equilibrio entre oferta y demanda evitando mermas provocadas por excedentes de producción.

2.4 Startups

2.4.1 Definición de *startups*

Podemos encontrar diferentes definiciones de *startup* cómo autores hablen de la materia. Intentos de definir el concepto que comenzaron a principios de la década de los 70's (primera aparición de los teléfonos celulares y ordenadores portátiles) y hasta en la actualidad no se ha conseguido llegar a una definición universalmente aceptada.

Se puede encuadrar a una *startup* cómo una empresa innovadora totalmente nueva en el mercado que presenta grandes posibilidades de crecimiento, es decir, modelos de negocios escalables en donde se detecta una oportunidad en determinado sector productivo y produce y/o comercializan un nuevo producto o servicio para satisfacer la necesidad detectada. Además, como lo indica su nombre “empresa emergente” el termino solo aplica en los comienzos del proyecto, cuando está iniciándose. Una vez instalada en el mercado ya no se definirá como *startup*. Es decir, las características más importantes de este tipo de organizaciones son: **temporalidad, crecimiento exponencial y escalabilidad.**

También, el término está muy vinculado en su origen con empresas con **base tecnológica**, son proyectos que apuntan una vinculación directa con la tecnología tanto como producto final, proceso o herramienta o como para proporcionar el producto/servicio final.

Los **factores de éxito** de este tipo de proyectos y su supervivencia depende, entre otras, de las siguientes variables según lo indicado por (Empresa-Concepto, 2012):

- Disponer de una tecnología o conocimiento maduro y consistente;
- Que la explotación de la ventaja competitiva que confiere el componente innovador de la tecnología o conocimiento esté amparada por una efectiva protección de la propiedad industrial, si procediese;

- Contar con personas con capacidad de liderar el proyecto y gestionar, a un grupo de personas con experiencia en I+D, altamente calificadas;
- Detectar a tiempo las oportunidades de mercado y orientar sus productos y servicios hacia los nichos existentes y
- Gozar de acceso a fuentes de financiación en todos los niveles: desarrollo de la tecnología o conocimiento, capital semilla, etc.

Se seleccionó a las compañías *startups* como caso de estudio ya que de alguna manera las nuevas empresas poseen la capacidad definir y diseñar propuestas de valor sin “ataduras”, lo que les da una amplia libertad a la hora de innovar en el modelo de negocio ofrecido.

2.4.2 Entorno actual en Argentina

Según un estudio realizado por *Global Entrepreneurship Monitor*¹⁵ (se realizó una muestra regional en Argentina de 4000 casos) arroja que el 15% de la población adulta tiene un emprendimiento o un negocio no mayor a tres años. Situación **que deja al país por encima del promedio mundial** del 12%, ubicando al país dentro del podio de los 20 países más emprendedores en el puesto número 16.

En línea con los datos enunciados anteriormente, se vislumbra un ecosistema alrededor del emprendedorismo argentino de ayuda que no existía años atrás, pero que ya estaba muy desarrollado en regiones como *Silicon Valley*, Los Ángeles, Tel Aviv, Singapur o San Pablo en el cuál abunda el capital del riesgo, la capacitación y la transferencia tecnológica y Buenos Aires había quedado delegada.

Con esto hacemos referencia al crecimiento y apoyo mediante la educación e inversión en actividades en las universidades que incorporan el apoyo al emprendedurismo en sus escuelas, organizaciones no gubernamentales, instituciones privadas y el gobierno nacional que por ejemplo con programas públicos y leyes buscan promover el financiamiento para el desarrollo y promoción de este tipo de negocios. Por ejemplo mediante **aceleradoras**¹⁶ o **incubadoras**¹⁷ cómo son los casos

¹⁵ *Global Entrepreneurship Monitor* (GEM): es un consorcio sin fines de lucro de investigación académica que tiene como objetivo hacer que la información de alta calidad sobre la actividad emprendedora global se encuentre fácilmente a disposición de un público tan amplio como sea posible. GEM es el estudio más grande de la actividad emprendedora en el mundo. Iniciado en 1999 con tan sólo 10 países, GEM ha llevado a cabo investigaciones en más de 80 economías de todo el mundo.

¹⁶ Aceleradoras: se enfocan en emprendimientos que se encuentran en una etapa más avanzada. Es decir, que ya cuentan con un modelo y un plan de negocios un poco más sólido, un producto ya definido y desarrollado, o en vías de desarrollo, y que tienen a sus equipos más establecidos. Más allá del capital extra, el espacio de *co-working*, y los *workshops* y la capacitación que suelen ser parte de la experiencia en una aceleradora, uno de los principales aportes que éstas hacen tiene que ver con las posibilidades de *networking*, y con la llegada a los inversores adecuados y a socios estratégicos que permitirán desarrollar el negocio y comenzar a escalarlo. (Pascual Fuste, 2017)

de *NXTPLabs* e *itBAF*, ambas dedicadas a los emprendimientos de base tecnológica. La primera de ellas se dedica a acelerar y financiar empresas de base digital que se encuentran en fase temprana y la segunda está abocado al negocio de los videojuegos y acompaña alrededor de 20 emprendimientos por año. También existen los denominados **inversores ángeles** que apadrinan empresas emergentes como El Club de Inversores Ángeles de IG o *Business Angels IAE*.

Si bien la actividad emprendedora aumentó en la última década, se trata de una proporción de emprendimientos domésticos bastante alta en Argentina si se tiene en cuenta, que muchos proyectos que surgen tienen que ver con la necesidad de ingresos debido al subempleo y la dificultad de encontrar trabajo, al igual que lo que sucede en los países menos desarrollados. En los países más avanzados, por el contrario, predominan proyectos más complejos en términos de uso de tecnología e innovación. Según el informe del G.E.M. las inestabilidades políticas y económicas (por sobre todo los procesos inflacionarios) traen como consecuencia que la población complete sus ingresos con una actividad de forma independiente. Pero además de la coyuntura nacional afirma que el crecimiento se da también en los emprendedores con una mejor “calidad” de proyectos con mayor potencial de impacto.

Por lo tanto, de acuerdo a lo anteriormente expuesto en este capítulo, este trabajo tiene el foco del estudio en la intersección de los siguientes tres tópicos desarrollados:

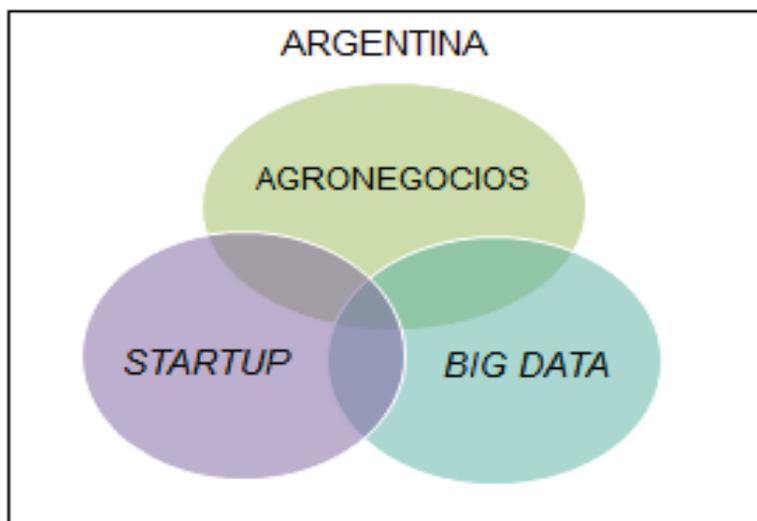


Imagen 5: Focos del trabajo. **Fuente:** Elaboración propia.

¹⁷ Incubadoras: Las incubadoras ayudan a los emprendedores que encaran esos proyectos a darle forma a sus ideas, a investigar el mercado, y hasta a entender quiénes son los potenciales clientes para cada producto. Para eso brindan herramientas como mentoría en áreas como emprendedurismo, finanzas, contabilidad, y en cuestiones legales, y ayudan al equipo a conectarse con otros emprendedores más experimentados, o que se encuentran en la misma fase, para ayudarlos a adquirir experiencia y conocimientos, y a definir cuestiones clave como su modelo de negocio. (Pascual Fuste, 2017)

Cada uno de ellos está relacionado en entender en qué contexto y cómo se están llevando adelante los proyectos de *startups* vinculados con los negocios agropecuarios en Argentina.

3. Big Data en startups agropecuarias

Para un productor agropecuario o ganadero es complicado crear la infraestructura necesaria para gestionar y analizar todos los datos generados en su actividad, pero actualmente hay empresas que han visto en éste sector un **nicho de negocio** importante para el desarrollo de aplicaciones basadas en *Big Data* fundamentado en los siguientes motivos. Primero, les permite realizar avances tecnológicos y aplicar sus conocimientos en un sector importante en la económica Argentina (con grandes posibilidades de implementar nuevos negocios por su capacidad de fondos). Por otro lado, la posibilidad de la gestión eficiente de los recursos naturales escasos para la preservación del medio ambiente y el reto que conlleva el aumento de la población con su demanda de alimentos exponencial que pone en riesgo la seguridad alimentaria genera todo un reto y desafío para la nueva generación de emprendedores para desarrollar herramientas disruptivas para un sector tradicional cómo lo es el agropecuario.

El sector está generando basta información digital debido a que sus actores, productores, contratistas, proveedores, utilizan nuevas maquinarias, aplicaciones y herramientas que generan datos que hoy se pueden digitalizar y ser procesados, analizados a altas velocidades gracias a la tecnología actual y es en donde entran en juego las nuevas *startups*.

Aquí se presenta un detalle de todas aquellas *startup* vinculadas al sector ganadero y al agropecuario que utilizan *Big Data* para el desarrollo de sus actividades principales que existen actualmente en el mercado argentino:

Empresa	Rubro	Función	Web site
FARMIN	Ganadero	Dispositivo capaz de pesar de manera individual y diaria cada cabeza de un rodeo vacuno	www.farmintech.com
CTM COLLAR	Ganadero	Collar que mediante la captación del movimiento de la vaca infiere en qué momento va a entrar en celo o si está sufriendo algún problema o enfermedad.	www.ctm-data.com
KILIMO	Agropecuario	Herramienta para el manejo de riego con el objetivo de recomendar estrategias de riego en función del consumo de cada cultivo.	www.kilimo.com.ar
BIOBOT	Agropecuario	Manejo del riego en agricultura extensiva. Sistema de controladores de riego conectados a internet y ajustan los tiempos de riego en base necesidades específicas.	www.biobot.com.ar
BOOSTER AGRO	Agropecuario	Aplicación que unifica todos los pronósticos de clima, en la cual el usuario localiza su campo en el mapa y puede ver lo que distintas fuentes están pronosticando.	www.boosteragro.com

SIMAPRO	Agropecuario	Monitorear cultivos y otras actividades como control de siembra, cosecha y aplicación de productos fitosanitarios.	www.monitoreoagricola.com
SISMAGRO	Agropecuario	Gestión agraria en la nube que permite llevar un registro preciso y actualizado en tiempo real de la actividad	www.sismagro.com
AGRO ZONE	Agropecuario	Permite sacar fotografías de los cultivos (soja, maíz, trigo) y obtener un <i>feedback</i> instantáneo de qué tipo de enfermedades y cuál es su grado de severidad	www.agrozone.ag
AGRO APP	Agropecuario	Sistema para predecir el clima y certificar Buenas Prácticas Agrícolas	www.agroapparg.com

Tabla 1: Resumen *startups* agro. **Fuente:** Elaboración propia.

A continuación se caracterizan las empresas descritas en la tabla precedente, en la cual se explica la metodología que utilizan para la implementación de *Big Data* cómo llevan adelante su modelo de negocios:

3.1 Sector ganadero

Farmin Tech:

- Ubicación geográfica: CABA
- CEO: Albornoz Ignacio
- Web Site: www.farmintech.com
- Logo:



Empresa incubada por el departamento de Investigación y Desarrollo de CREA desarrollaron un **dispositivo capaz de pesar de manera individual y diaria cada cabeza de un rodeo vacuno** sin movilizar ni embretar la tropa. Funciona en base a una plataforma de pesaje que se coloca en los bebederos y una antena que lee la información del chip que tienen las caravanas de identificación. Cuando uno de los animales o varios de ellos se suben a la plataforma con las cuatros patas encima de ella, la antena lee la caravana y a través de un algoritmo se van individualizando los datos para tener información en tiempo real del pesaje de los animales. Entonces, **basada en tecnología IoT la empresa desarrolló una red para transmitir datos en el campo a grandes distancias**, con diferentes sensores (desde monitores de rendimiento en cosechadoras hasta humedad del suelo o comportamiento animal). Por ejemplo el dispositivo ganadero genera información precisa sobre qué animales ganan más o menos peso, además de saber si están yendo o no a la aguada. Esto permite

tener un mayor control para el manejo de la hacienda, ya sea para dar por terminado un conjunto o, por ejemplo, cambiarle la dieta.

CTM Collar

- Ubicación geográfica: CABA
- CEO: Guido Buscetti
- Web Site: www.ctm-data.com
- Logo:



CTM Data (*Cattle Tracking Monitor*) desarrollaron un **collar que mediante la captación del movimiento de la vaca infiere en qué momento va a entrar en celo o si está sufriendo algún problema o enfermedad**. Captan además cuántos pasos da por día, a cada hora, y la inclinación de su cuello y luego reporta a una base que luego envía las estadísticas a la web. Utilizan **tecnología de radiofrecuencia (LoRa¹⁸)** de largo alcance y bajo consumo que permite tener un dispositivo que puede comunicarse con la base central estando hasta a tres kilómetros de distancia. Y en los casos más extremos, donde se requiere de una mayor cobertura, realizaron una alianza con SIGFOX, una empresa francesa que brinda **conectividad para dispositivos IoT**. La otra rama del negocio es la **venta de los datos** que la empresa va recolectando. En cuanto a la confidencialidad de los datos afirma el CEO Guido Buscetti que realizan un contrato con los productores en la cual informa que se mantendrán la información de forma confidencial, sin especificar ningún tipo de dato personal, con el fin de obtener parámetros estadísticos que permitan la mejora del sistema.

3.2 Sector agropecuario

Kilimo

- Ubicación geográfica: Córdoba
- COO: Rodrigo Tissera
- Web Site: www.kilimo.com.ar

¹⁸ LoRa: tecnología para conexiones a grandes distancias y para redes de IoT que se pueden utilizar en ciudades inteligentes, lugares con poca cobertura celular o redes privadas de sensores o actuadores. Entre sus principales ventajas se encuentran: alta tolerancia a las interferencias, alta sensibilidad para recibir datos, largo alcance, baja transferencia de datos.

- Logo:



Es una aplicación que por medio de los **análisis de suelo se conoce el contenido hídrico**, es decir el estado inicial de la cuenta y luego a través de datos satelitales y de estaciones meteorológicas de acceso libre se modela la evapotranspiración¹⁹, el gasto de la cuenta y luego al productor le indica cuánto llovió y regó, **con el objetivo de recomendar estrategias de riego en función del consumo de cada cultivo**, con el fin de hacer más eficiente el uso del recurso natural y mejorar los rindes.

Kilimo ganó el primer programa *Agrotech* de Latinoamérica, organizado por la **incubadora NXP Labs**, que le permitió su primera expansión. Además luego ganaron el **programa AgLaunch**²⁰ lo que le permitió acceder a una inversión inicial de 50 mil dólares para montar su proyecto en Estados Unidos, además de asesoramiento, la posibilidad de llegar a una amplia red de productores de esa región, y sitios de validación en el área. La aplicación cerró una nueva ronda de capital semilla por 850 mil dólares liderada por los fondos de inversión *Xpand Ventures* y *Alaya Capital Partners*.

Biobot

- Ubicación geográfica:
- CEO: Santiago Fux
- Web Site: www.biobot.com.ar
- Logo:



Desarrollaron un **sistema de controladores de riego** conectados a internet que **ajustan los tiempos de riego en base necesidades específicas**. Producen mejoras en la calidad y productividad del cultivo, y disminuye el mantenimiento preventivo, en cuanto a las horas de trabajo y materiales de reposición se refieren. Obtiene información del clima a través de internet y en conjunto con sensores colocados *in situ*

¹⁹ Evapotranspiración: Cantidad de agua del suelo que vuelve a la atmósfera como consecuencia de la evaporación y de la transpiración de las plantas.

²⁰ AgLaunch: es un programa de 15 semanas enfocado en la agricultura que incluye un currículo de clase mundial y mentoría. Establecido en Memphis, el programa está anclado en el acceso a la producción agrícola a través de múltiples cultivos y entornos de producción a través de relaciones con agricultores y compañías de toda la región.

(sensores de caudal, humedad, temperatura ambiente y humedad del suelo, etc.) se puede determinar cuánta agua se debe aplicar para lograr un balance hídrico constante en el cultivo para potenciar el crecimiento del mismo. Es decir, **utilizan información de dispositivos implantados en el sector (IoT) y de fuentes externos de la web** (cómo por el ejemplo el servicio meteorológico de la zona geográfica en la que está implantado el dispositivo) para alimentar de esta forma los resultados del análisis.

El usuario del servicio puede acceder a través de la aplicación y **monitorear a través de tablero de control y visualizar que es lo que está ocurriendo en tiempo real**. Almacenan en las bases de datos minuto a minuto los datos generados por cada sensor, permitiendo que los usuarios puedan visualizar esos valores a través de por ejemplo la cantidad de litros acumulados detectados por los caudalímetros. Hoy en día su CEO asegura que ofrecen una actualización de los datos cada hora ya que es más útil para el usuario su análisis y procesamiento. Cada usuario tiene sus controladores y en función de eso, se garantiza por privacidad y confidencialidad de los datos de caudal, de humedad, etc. Pueden acceder a los reportes, por ejemplo, inclusive exportar esos datos, pero son privados.

Booster Agro

- Ubicación geográfica: CABA
- CEO: Marcos Alvarado
- Web Site: www.boosteragro.com
- Logo:



Es una aplicación (disponible gratuitamente) que reúne información disponible y dispersa en la web para agruparla y facilitar la tarea al usuario **unificando todos los pronósticos de clima**, en la cual el usuario geolocaliza su campo en el mapa y puede ver lo que distintas fuentes meteorológicas pronostican para un área geográfica específica. Usualmente los productores agrícolas consultan a diario varias fuentes de información ya que la seguridad de los pronósticos es baja y es un factor clave para la gestión de sus tierras. Además, puede consultar un registro histórico de temperaturas medias, lluvias acumuladas zonales, sincronizarlo con su planilla de lluvias y visualizar de forma remota las condiciones actuales de su campo, sin necesidad de instalar algún *hardware* en el campo. Otra vertiente del servicio **es resguardar la información**. Esta es una ventaja que ofrece, ya que los datos de los productores son anónimos y están protegidos bajo normas de seguridad. La aplicación es utilizada por productores, administraciones y grandes explotaciones. La aplicación ayuda a

centralizar toda la información en un mismo lugar, para una gestión más ágil y ahorrando tiempo. Para solucionar los problemas de conectividad que pueden acarrear de las zonas geográficas más alejadas y suplantar la cobertura de red y el acceso a internet utilizan para proveer la información canales como mensajes de texto (SMS), operadoras telefónicas (IVR) y servicios de mensajería instantánea (USSD), consideraron fundamental poder operar con canales de comunicación que estén adaptados al contexto.

Simapro

- Ubicación geográfica: CABA
- CEO: Santiago Arias
- Web Site: www.monitoreoagricola.com
- Logo:



Es una *startup* que desarrolló una aplicación para teléfonos y *tablets* que permite a productores, ingenieros agrónomos y empresas de producción del agro **monitorear cultivos** y otras actividades como **control de siembra, cosecha y aplicación de productos fitosanitarios**. El objetivo es que el proceso sea menos tedioso y, a la misma vez, digitalizar todos los datos en origen, de manera que no se pierdan y enviar luego un análisis sobre los mismos. El producto desarrollado usa las funcionalidades de los *smartphones* como GPS, cámara de fotos, micrófono, etcétera. Que luego todos estos datos pueden ser luego examinados en detalles en una plataforma web en la cual el productor crea un usuario y obtiene los reportes en conjunto con los tableros de control para la toma de decisiones.

Sismagro

- Ubicación geográfica: CABA
- CEO: Christian Zechner
- Web Site: www.sismagro.com
- Logo:



Ofrece una plataforma de **gestión agraria en la nube que permite llevar un registro preciso y actualizado en tiempo real de la actividad**. Los productores crean una cuenta y dibujan los lotes de su campo en el mapa. Luego pueden registrar las labores realizadas en esos lotes para por ejemplo sacar el presupuesto de cada actividad

realizada en ese lote, el valor de los insumos consumidos para llevar a cabo determinada actividad (fertilizantes, mano de obra, semillas, etc.), controlar los stocks de los galpones, ver el rendimiento de cada cultivo y además visualizar la posible ganancia de cada lote. De esta forma se genera un historial de lo sucedido en el campo para optimizar la toma de decisiones.

Agro Zone

- Ubicación geográfica: CABA
- CEO: Fernando Derossi
- Web Site: www.agrozone.ag
- Logo:



Es una *startup* nacida en el MIT (Instituto Tecnológico de Massachusetts) con la visión de **maximizar la eficiencia en producción de alimento** a través de tecnología y ciencia de datos aplicada a la agricultura. La *App* permite sacar fotografías de los cultivos (soja, maíz, trigo) y obtener un *feedback* instantáneo a través de un algoritmo de qué tipo de enfermedades y cuál es su grado de severidad. De esta manera, realiza un diagnóstico y luego, mediante un algoritmo, recomienda los productos a utilizar. También tiene una solución para plagas y malezas, que incrementa el rendimiento del productor mediante la reducción de pérdidas en el cultivo. Esto se realiza a través de técnicas de **Machine Learning**, por eso gracias a esta tecnología, cuando un productor sube una fotografía o una información al portal, la hace cada vez más inteligente ya que a través de un código que va enseñándole y eleva el nivel de precisión, es por esto que utiliza bases de datos voluminosas, a las que se le añaden interfaces de información sobre el clima, tipo de suelo, temperatura, condiciones predisponentes para enfermedades, etc. El algoritmo utilizado para la recomendación de productos se basa primero en la eficiencia, basado en investigación (académica y de campo); el segundo, la sustentabilidad. Es decir, se recomienda, el que sea más efectivo o “banda verde”. En caso de haber dos productos en conflicto (porque ambos son aplicables) se analiza la cantidad de ingredientes activos por hectárea que utiliza. Además de la identificación, la *App* analiza el porcentaje de severidad y ofrece una recomendación de tratamiento, teniendo en cuenta qué productos y qué distribuidores hay disponibles en la zona

Al inicio, el productor carga solamente un número de hectáreas, genética que sembró y algunos otros datos básicos. Pero, luego va agregando datos con cada foto que saca queda geo-localizada, con horario, día, condiciones de temperatura, humedad, etc. Con esa información, más el aporte que se le agrega sobre qué pasa en diferentes

zonas del país, cómo vienen los vientos y otros sets de información, se puede predecir y saber anticipadamente qué va a pasar.

El *software* está diseñado para productores agropecuarios que tienen descarga gratuita y reciben ingresos no a través de la descarga sino **por la venta de mapas e información** pertinente al sector medida por la aplicación a diferentes actores del resto de la cadena que pueden ser distribuidores, compañías financieras, empresas de agroquímicos que usen la información de manera agregada, con comparaciones y mapas, es decir, obtiene sus ingresos a través de la venta de datos que obtiene de los productores. Es por esto, que para este grupo de público la App es gratuita ya que se nutren de información de ellos dándole como recompensa información útil para el manejo de los lotes sin costo alguno para el productor.

AgroApp

- Ubicación geográfica: CABA
- CEO: Esteban Videla Pearson
- Web Site: www.agroapparg.com
- Logo:



En el marco de un proyecto del FONTAR, programa del Ministerio de Ciencia, Tecnología e Innovación Productiva de la Argentina (Mincyt), se creó un sistema para **predecir el clima y certificar Buenas Prácticas Agrícolas**²¹. Se basa en **clima de precisión, internet de las cosas y algoritmos de inteligencia artificial** mediante una red de más de 100 estaciones conectadas a internet. Además, ofrecen un servicio de pronóstico climático, que *rankean* entre aquellos que mejor se adaptan a cada región.

Se utiliza inteligencia artificial, ya que se mide el dato que extraen de la estación con el mejor de los pronósticos, que luego se mejora mediante algoritmos de aprendizaje automático para mejorar la precisión en la información brindada al productor para generar un sistema que sea confiable y seguro. La misma aplicación también hace **recomendaciones automáticas** en base a las condiciones meteorológicas que está censando y midiendo constantemente, por ejemplo si las

²¹ Buenas Prácticas Agrícolas: Son un conjunto de principios, normas y recomendaciones técnicas aplicables a la producción, procesamiento y transporte de alimentos, orientadas a asegurar la protección de la higiene, la salud humana y el medio ambiente, mediante métodos ecológicamente seguros, higiénicamente aceptables y económicamente factibles. La gestión responsable de fitosanitarios tiene como objetivo lograr el manejo y uso responsable de los fitosanitarios durante todo su ciclo de vida: desde su descubrimiento y desarrollo, ciclo comercial y uso en el campo, hasta su eliminación por el uso y disposición final de envases.

temperaturas son elevadas y no hay pronóstico de lluvia en los días cercanos, probablemente le recomiende utilizar algún producto antievaporante²².

Cuentan además con un Plan de Municipios Sostenibles que facilita datos e inteligencia aplicada a la comunidad del municipio con el objetivo de que se logren buenas prácticas agrícolas mediante auditoría colectiva, construyendo así una licencia social. Esta licencia mantiene la eficiencia de la aplicación fitosanitaria sin dañar el ambiente, los municipios sostenibles son:

- Balcarce;
- Pergamino;
- Junín;
- Lincoln;
- San Antonio de Areco y
- Olavarría.

²² Producto antievaporante: Es un aceite vegetal producido a partir de aceite de soja, fácilmente emulsionable en agua que reduce la evaporación y tensión superficial de los productos aplicados, mejorando la calidad de aplicación.

4. Resultados y análisis

Con el fin de tener una perspectiva más certera sobre cómo se están configurando los modelos de negocios basado en la tecnología de datos y con el objetivo de encontrar parámetros similares entre las distintas estructuras de organización, como se detalló al inicio del presente trabajo, se entregaron a las distintas *startups* (aquellas mencionadas en el apartado anterior) un cuestionario de diez preguntas con opciones múltiples para recabar información para el presente trabajo que no se encuentra disponible a través de otras fuentes que no sea propia de cada integrante de la organización. Las personas encuestadas forman parte de la gerencia de la empresa, entre ellos primordialmente son CEOs y/o co-fundadores o cargos pertenecientes a las áreas de operaciones o al área comercial. El cuestionario se envió a través de un correo electrónico en el cual se formalizó el pedido y la presentación de las preguntas que ellos mismos respondieron y avalaron la información volcada a través de una respuesta confirmatoria.

Antes de entrar por completo a la presentación de los resultados encontrados, cabe mencionar que en términos generales las empresas tienen de desarrollo entre un año y medio y dos años de pleno funcionamiento, en las cuales el número de integrantes participantes varía entre las tres y cinco personas con un rango etario por debajo de los treinta y cinco años de edad cada uno. Además se evidenció que los equipos que llevan adelante estas organizaciones son interdisciplinarios, es decir, que son conformados por distintas profesiones, no solo aquellas enfocadas en la gestión o el manejo de datos. Entre ellos las más preponderantes son las vinculadas al sector agropecuario como los ingenieros agrónomos o veterinarios, pero también conforman parte de los equipos ingenieros en sistemas y/o electrónicos (especialistas en el desarrollo del *software* y *hardware*), físicos y licenciados en administración de empresas entre otros. Otra arista importante para destacar es que provienen en su gran mayoría del rubro agropecuario, tanto sea porque trabajan en grandes compañías con anterioridad (por ejemplo Monsanto, Basf o Nidera) o también por herencia familiar, es decir, que sus padres o abuelos se dedicaron al rubro y ellos heredaron la profesión familiar.

El primero de los parámetros indagados fue sobre las **barreras de entrada** más importantes que encuentran para la implantación de empresas vinculadas al *Big Data*. Es decir, aquellos obstáculos que complicaron o dificultaron el desarrollo de cada organización. El 30% de las empresas encuestadas sugiere que la falta de recursos humanos especializados disponibles constituye una barrera de ingreso, al igual que la mitad de los encuestados (el 50%) declara que la falta de criterio sobre la rentabilidad

(ROI)²³ que puede ofrecer este tipo de inversiones es un obstáculo a superar. Se sostiene que aún al ser la temática novedosa y no tener registros con antelación de este tipo de proyectos es dificultosa la selección de personal adecuado que contenga las capacidades técnicas requeridas y al igual que la medición clara de un objetivo de rentabilidad para estas empresas. Se pueden realizar aproximaciones y estimaciones de cómo va a reaccionar el mercado o cuáles serán las proyecciones de ventas estimadas pero son variables de gran incertidumbre a la hora de comenzar a decidir la creación de estas organizaciones, ya que no hay antecedentes claros de cuál puede ser la reacción o el efecto de las políticas de negocio llevadas a cabo.

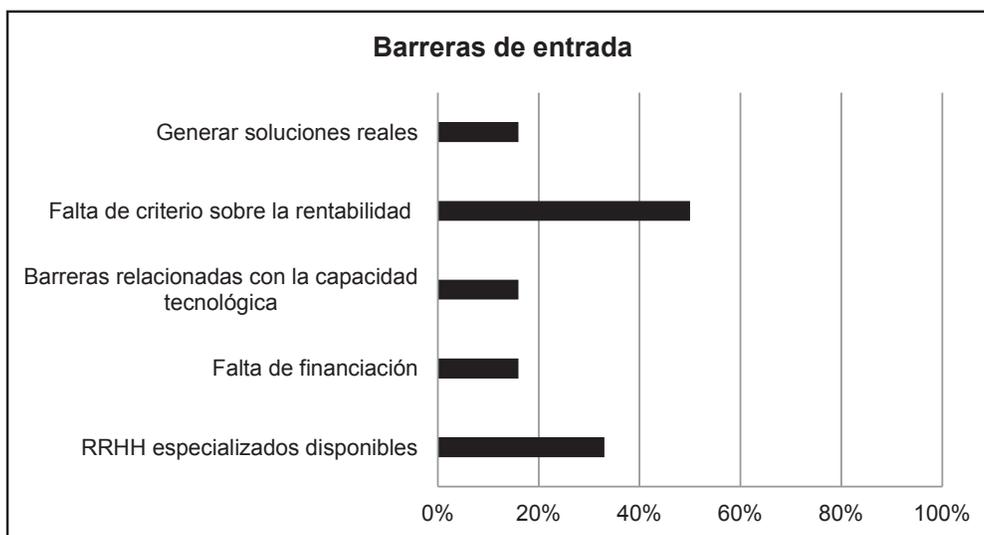


Gráfico 4: Barreras de entrada. **Fuente:** Encuesta *startup* agro *Big Data* 2018

Se averiguó también sobre los **factores limitantes que encuentran los fundadores de las empresas al momento de vender los servicios** al cliente, cuales son aquellas trabas que impedían la finalización de las ventas.

Todas las empresas encuestadas señalan que los gastos de instalación y mantenimiento son elevados para que los productores puedan añadirlo a su matriz de costos, justificando que ya cuenta con una carga elevada debida a los aumentos de energía eléctrica, los costos logísticos y la alta presión tributaria. Cómo se describió en el punto anterior generalmente las empresas de *Big Data* requieren de instalaciones de *hardwares* para la recolección de datos más la información brindada a través del *software* y el monto mensual del mantenimiento, implicando un costo adicional a la hora de ofrecer el servicio a los productores. Adicionado a lo descripto anteriormente la generación de agricultores y ganaderos presentes hoy en el campo argentino aún no está capacitado en su gran mayoría para lidiar con el manejo de un *software* que contenga todos los datos integrados y con su equipamiento. Es por esto, que la falta

²³ ROI (*Return On Investment*): Es el valor económico generado como resultado de la implementación de diferentes acciones. Este indicador permite medir el rendimiento obtenido de una inversión

de entendimiento por parte del cliente figura cómo la opción número dos en cuanto a los factores limitantes.

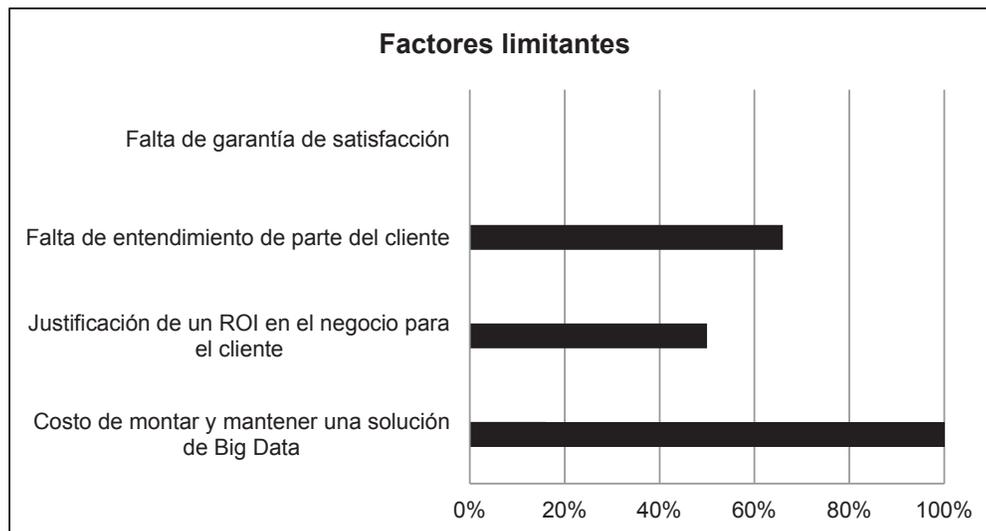


Gráfico 5: Factores limitantes. **Fuente:** Encuesta *startup* agro *Big Data* 2018

Para mitigar con los inconvenientes de entendimiento de parte del cliente, los CEOs se vieron obligados a desarrollar políticas de asistencia personalizada para una **relación con su segmento objetivo más satisfactoria**. Dando como resultado un 100% de selección en la respuesta “asistencia personalizada” sobre el resto de las opciones disponibles.

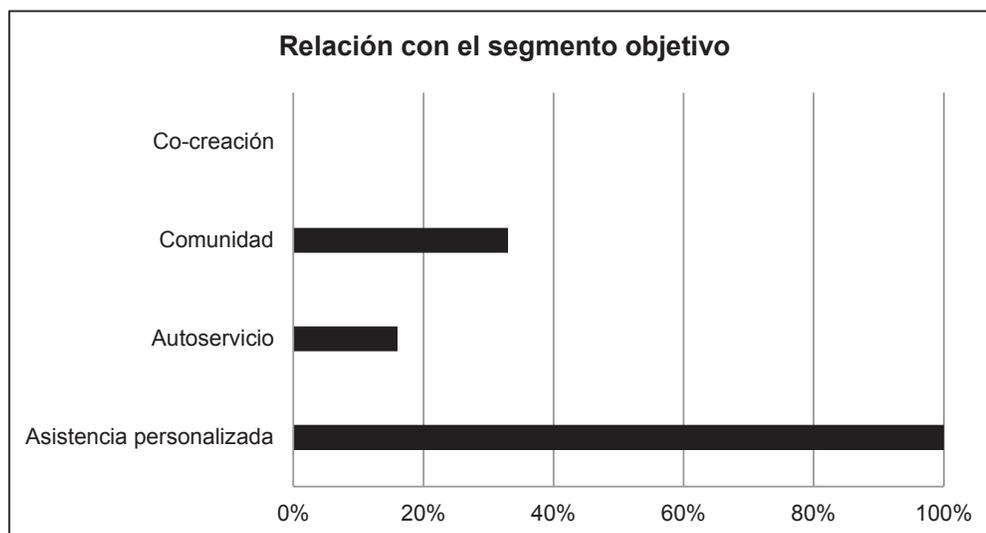


Gráfico 6: Relación con el segmento objetivo. **Fuente:** Encuesta *startup* agro *Big Data* 2018

En cuanto al **tipo de cliente actual o potencial** al que se enfrentan el 66,7% indicaron que pertenecen al grupo de clientes B2B²⁴. Esto quiere decir que las aplicaciones son

²⁴ B2B (*business to business*): hace referencia a las transacciones comerciales entre empresas.

utilizadas por ejemplo por productores chicos, medianos, administraciones agropecuarias y/o grandes explotaciones agrícolas.

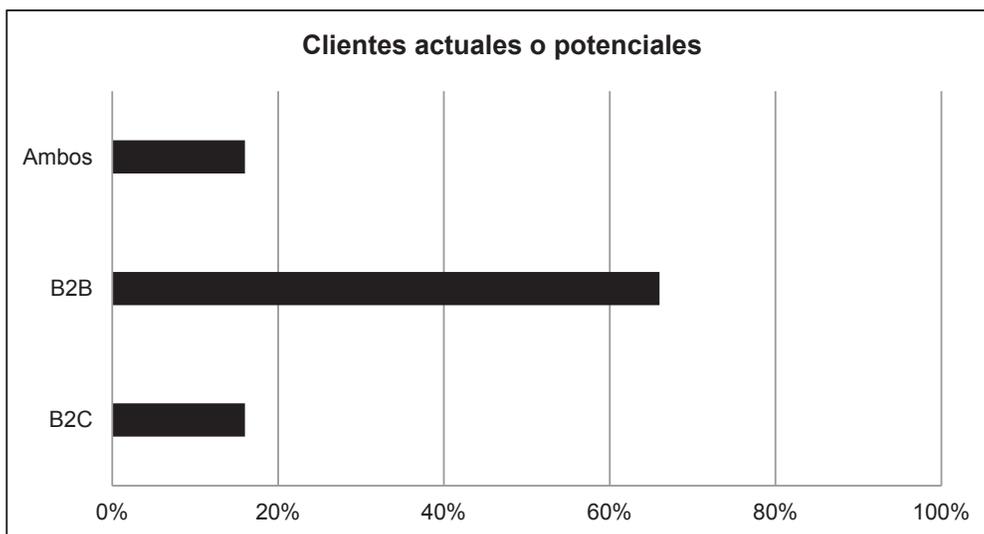


Gráfico 7: Tipo de clientes. **Fuente:** Encuesta *startup* agro *Big Data* 2018

La encuesta también hace referencia sobre **los tipos de fuentes de información que utilizan los empresarios para el desarrollo de los servicios**. Entre las más seleccionadas se encuentran las provisionadas a través de los propios clientes o internas. Ambas provenientes del método M2M, como por ejemplo en el caso de CTM Collar con el collar colocado a las vacas o en Farmin con la plataforma de pesaje transmiten información proveniente del cliente para “alimentar” los reportes de salida o también los sensores de Biobot que transmiten información respectiva al caudal, humedad, temperatura ambiente y humedad del suelo.

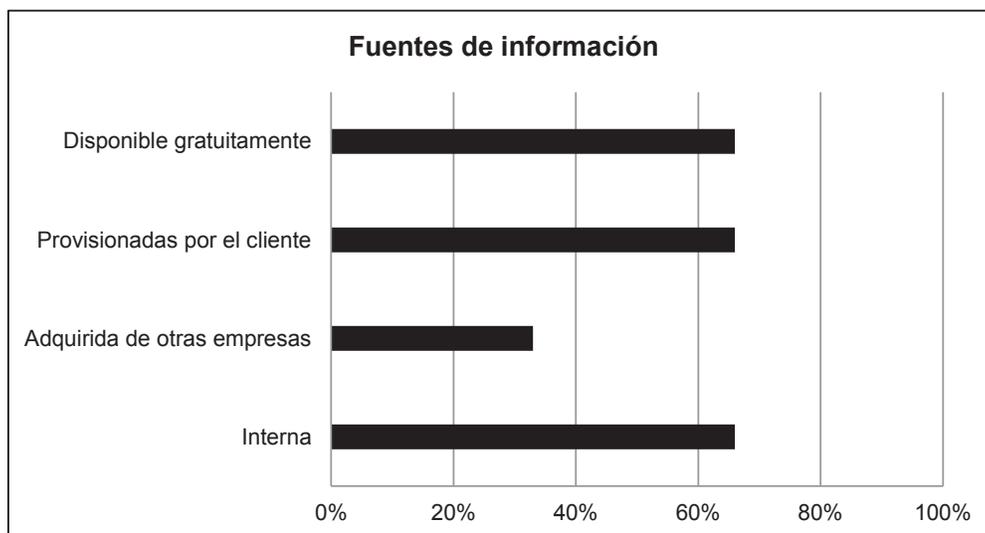


Gráfico 8: Fuentes de información. **Fuente:** Encuesta *startup* agro *Big Data* 2018

Además existen el caso por ejemplo de Simapro en el cual el productor es quien además nutre de datos al sistema a través de la registraci3n de malezas, plagas y

enfermedades simplemente tomando fotografías, notas de voz o texto a medida que va recorriendo los lotes. Este tipo de empresas también podemos notar que se nutren de fuentes disponibles gratuitamente en los servicios web, en su mayoría utilizan información referente servicios meteorológicos.

Otras de las vertientes preguntadas en el cuestionario era para comprender según la visión de los fundadores cuales creen que en base a sus criterios son **los diferenciales del negocio**. Entre las opciones propuestas las más destacadas con el 50% de elección entre los encuestados, fue la capacidad de las empresas de conjugar datos heterogéneos y la capacidad también de integrar datos de distintas fuentes con la misma cantidad de elección entre los encuestados. Esto nos indica que trabajan en base a las cuatro V de *Big Data* mencionadas en el capítulo número 2. En cuanto a la utilización de inteligencia artificial podemos notar que no es una metodología extendida dentro de estas organizaciones. El caso que más se destaca entre el universo de empresas es el de AgroAPP que utiliza IA para que a través de diferentes algoritmos de aprendizaje automático lograr la mejor precisión en la información brindada al productor para que el sistema sea confiable y seguro.

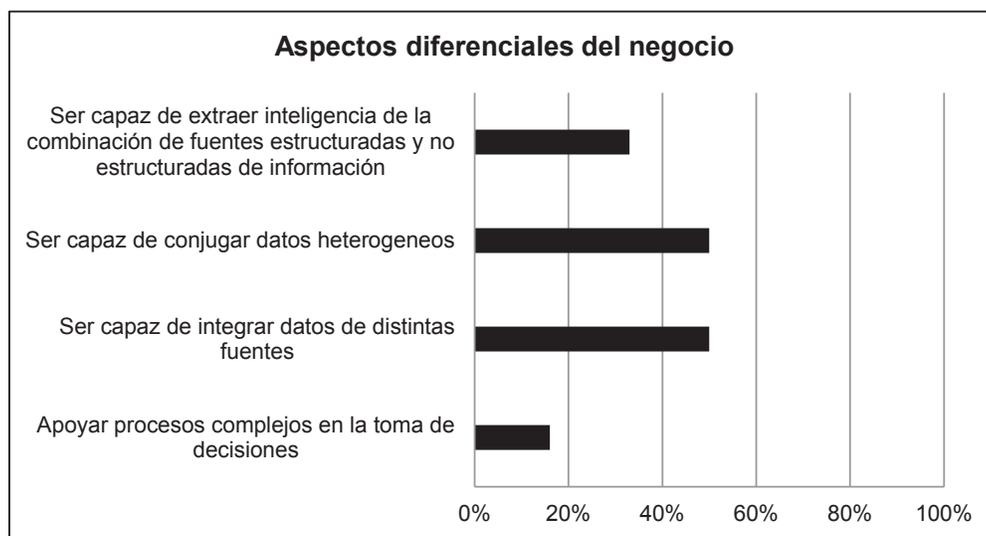


Gráfico 9: Aspectos diferenciales del negocio. **Fuente:** Encuesta *startup* agro *Big Data* 2018

En correlación con el gráfico anterior y en línea con el concepto de *Big Data* **las empresas seleccionaron como actividad clave** en su mayoría (83%) la visualización, análisis y procesamiento de los datos. La visualización de la información la realizan a través de las aplicaciones con tableros de control o recomendaciones automáticas a los productores y en conjunto con el análisis y el procesamiento de los datos utilizan técnicas como el *machine learning* o inteligencia artificial, ambas seteadas con algoritmos propios de cada empresa necesarios para por ejemplo para Booster Agro procesar y entregar 15.000 datos climáticos por hora y más de 10 millones por mes. En cuanto a la adquisición de datos, éstas empresas utilizan un

hardware como el caso de CTM Collar que ofrecen un servicio de DaaS²⁵ (*Data as a Service*) mediante un dispositivo (el collar colocado a las vacas).

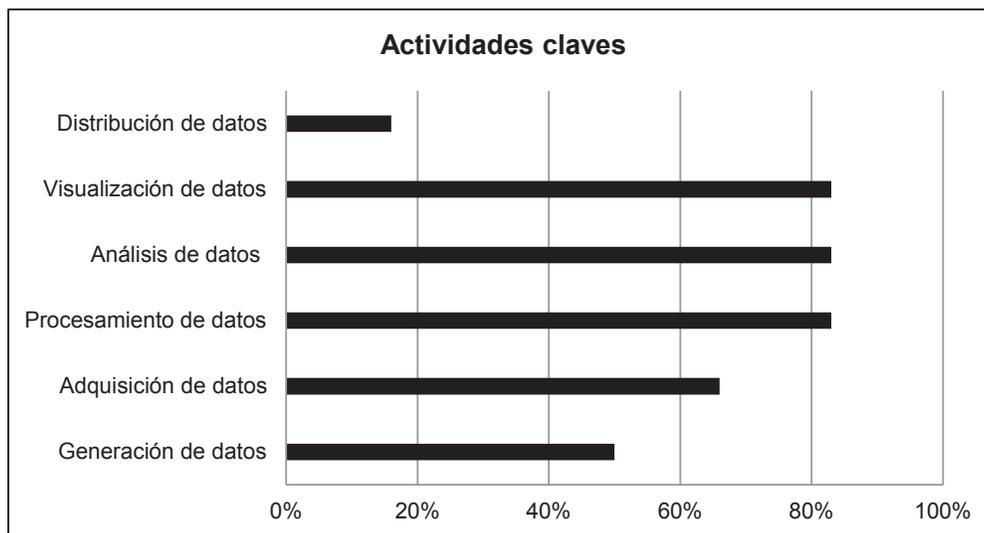


Gráfico 10: Actividades claves del negocio. Fuente: Encuesta *startup* agro *Big Data* 2018

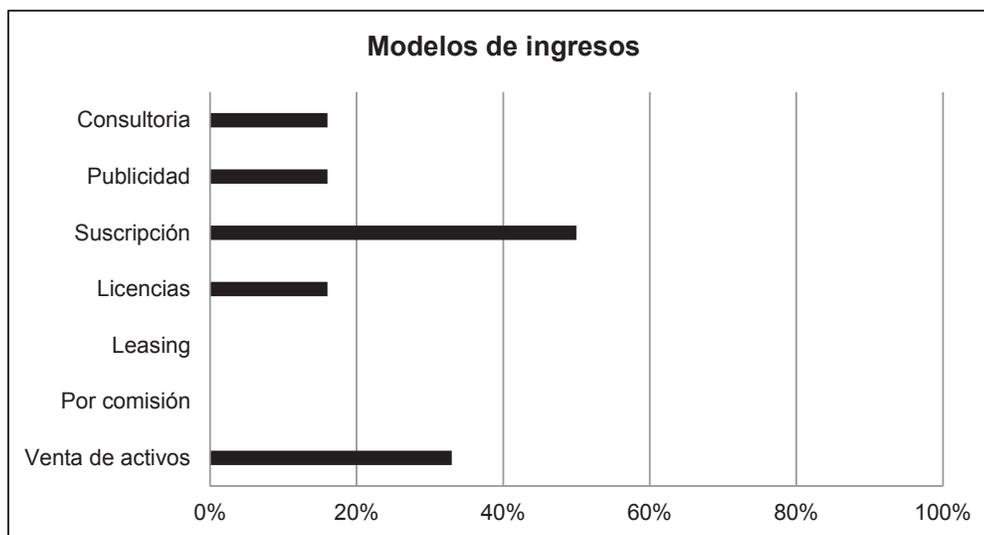


Gráfico 11: Modelos de ingresos. Fuente: Encuesta *startup* agro *Big Data* 2018

El parámetro número ocho evaluado fue el sobre **el modelo de ingresos que lleva adelante cada empresas**. Cómo lo indica el gráfico precedente a estas líneas el 50% de las empresas seleccionaron cómo opción principal la suscripción, seguidamente con un 40% de la venta de activos y con 16% el asesoramiento a través de consultorías y publicidad directa en sus sitios web o aplicaciones móviles. En varios casos la descarga y utilización de la aplicación para los productores es gratuita, cómo

²⁵ DaaS: es un modelo de provisión y distribución de la información en la cual los archivos de datos (incluyendo textos, imágenes, sonidos y videos) se ponen a disposición de los clientes a través de una red, normalmente internet.

es el caso de Agro App, Sismagro, Agrozone y Booster Agro. Pero cada uno administra de forma diferente sus cuentas. En el caso de Agrozone, su principal negocio es vender a los proveedores de productos agrarios información que les permita adecuar sus existencias a las necesidades locales, estos proveedores pueden ser distribuidores, compañías financieras, empresas de agroquímicos que usen la información de manera agregada. Además tienen la posibilidad de que marcas puedan publicitar y dar a conocer sus productos o servicios a través de una sección en la aplicación que obtiene resultados únicamente patrocinados. Pero por ejemplo para Booster Agro el caso es diferente, ellos cuentan con una versión avanzada de la aplicación que no es gratuita y contiene mayor tipo de información y *reporting*. Para Sismagro cuenta con una versión gratuita por tiempo ilimitado, pero con restricciones cómo administrar un solo usuario, un campo, 10 lotes y una determinada cantidad de superficie según la actividad. La versión customizada tiene un costo a partir de 19\$US por mes. Sismagro por ejemplo cobra el servicio por cantidad de lotes cargados al sistema un costo promedio de US\$ 0,20 por hectárea que se renueva en forma anual. Este precio varía de acuerdo a la cantidad de hectáreas totales contratadas. A más hectáreas, menor precio

Con respecto a la venta de activos físicos, también CTM Collar entrega los dispositivos en comodato (no se venden los collares) y se paga un *fee* mensual de US\$ 5 por animal a cambio de la información y la otra variante del negocio es la venta de datos recolectados a distintos actores del mercado. Pero por ejemplo para BioBot el *hardware* se vende a un precio de US\$ 300 y por su parte, la plataforma tiene un plan gratuito para los clientes residenciales y un abono de US\$ 150 mensuales para uso profesional.

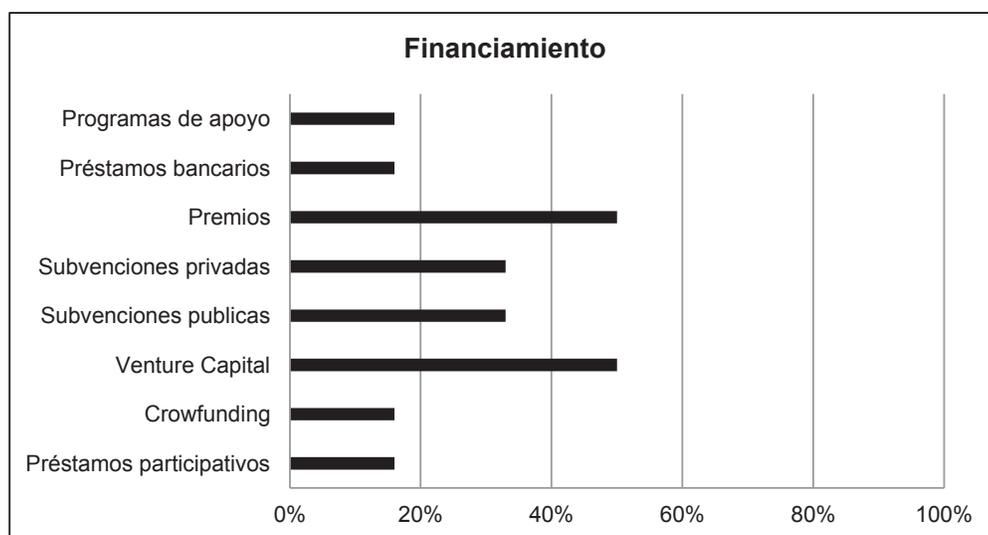


Gráfico 12: Financiamiento. Fuente: Encuesta *startup* agro *Big Data* 2018

En cuanto al **parámetro de financiamiento**, podemos observar que la mitad de los encuestados recibió financiamiento por parte del modelo *venture capital* y a través de

premios. Seguido con el 30% las empresas tuvieron acceso a tanto subvenciones privadas como públicas entre las opciones más destacadas. Podemos mencionar el caso de Farmin que recibió premios para su desarrollo de parte de *Crea Tech* y *AG Innovation Showcase*, por el lado de CTM Collar recibieron un subsidio por parte del Fonsoft (Fondo Fiduciario de Promoción de la Industria del Software) del Ministerio de Ciencia y Tecnología de la Nación al igual que BioBot que percibió un aporte no reembolsable de 10 mil dólares también del Fonsoft. Kilimo por su parte accedió al programa *AgLaunch* que le permitió obtener una inversión inicial de 50 mil dólares. Además de recibir el apoyo de la aceleradora Tennessee, Kilimo también ganó el primer programa Agrotech de Latinoamérica, organizado por NXTP Labs. En el 2018 obtuvo ronda de inversión que le significó una capitalización de 850 mil dólares, con aportes de Xpand Ventures y Alaya Capital Partners. Las dos aceleradoras inyectaron capital a Kilimo en cuotas similares a la Agencia Córdoba Innovar del Gobierno provincial. Este aporte se dio en el marco del programa Córdoba Acelera, del que Xpand y Alaya son beneficiarias: por cada peso aportado por las privadas, el Estado puso otro. Así, la Agencia redondeó una inversión de 150 mil dólares en el emprendimiento, una apuesta de *venture capital* sin precedentes en el sector público. Esto nos indica que se comienzan a vislumbrar la intervención del Estado a través de subvenciones y apoyos financieros en mercados nuevos como el generado por *Big Data* y continua marcando una diferencia con la encuesta enunciada en el apartado 2.2 realiza por el CIECTI en la cual indica que la intervención del Estado en proyectos de *Big Data* es escasa. El equipo de Biobot por ejemplo tiene acumulado los siguientes premios:

- BA IoT
- Crea
- Campus Emprear
- *Startup* Oeste
- Red Innova Challenge
- Programa Agrotech NXTP
- Alltec+100
- BID Idear Soluciones

En cuanto a la **clase de relaciones que desarrollaron las empresas para poder lanzarse al mercado** podemos notar una paridad en la mayoría de las respuestas con un nivel de selección por encima del 60% entre aceleradoras, incubadoras, centros de investigación e instituciones de ciencia y tecnología. Esta elección realizada indica que todas las empresas han optado por la vía “habitual” que siguen las *startups* para el desarrollo y crecimiento.

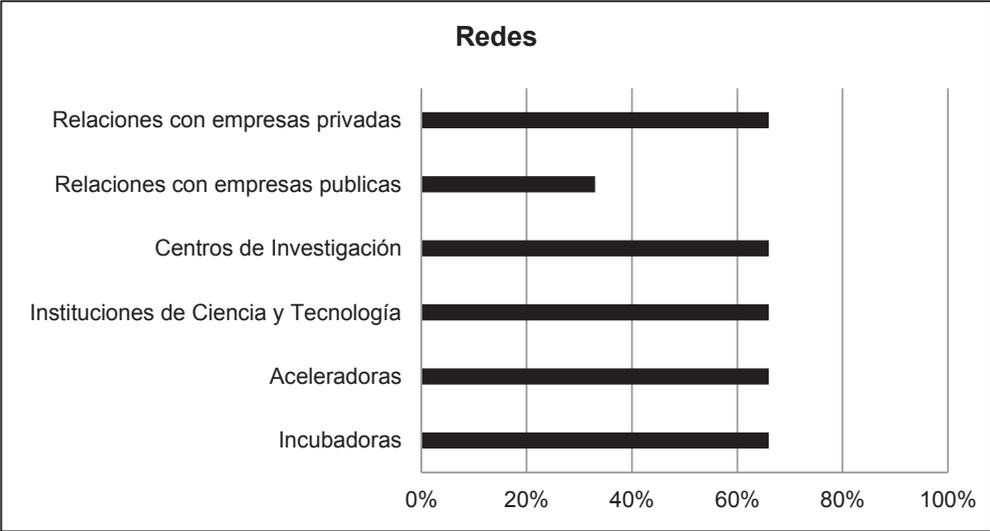


Gráfico 13: Redes. **Fuente:** Encuesta *startup* agro *Big Data* 2018

Podemos referir en base a los resultados presentados que las *startups* mencionadas se encuentran en un proceso de crecimiento, por lo tanto van perfeccionando sus métodos y adaptándose a los nuevos requerimientos de parte de los clientes, mejorando sus propuestas de valor para hacerlas cada vez más atractivas para su segmento objetivo. Para lograr convertirse en empresas sostenibles en el largo plazo deben acentuar la actividad la gestión de los datos y en el asesoramiento a los clientes para ir generando una cultura en el uso de los datos y que sea un aliado de sus negocios y no una carga de costos en sus estructuras, intentando influir cada vez más en la toma de decisiones

5. Conclusiones

Durante el recorrido de éste trabajo se puede observar que la gestión de los datos masivos ha empezado hace ya unos años a cambiar la forma de hacer los negocios, los mercados y la sociedad, convirtiéndose en un activo corporativo importante. Es decir, quien tenga éste recurso a disposición podrá obtener una ventaja comparativa frente al resto de sus competidores. Éste camino comenzó de la mano de las grandes compañías, quienes fueron pioneras en las técnicas de procesamiento de datos masivos. Pero actualmente se vislumbra una democratización de las técnicas y metodologías y gracias a esto podemos notar el florecimiento de nuevas empresas que se insertan en el mercado de *Big Data*.

Estas nuevas empresas que son lideradas por jóvenes profesionales generan un cambio disruptivo en los agronegocios y en la forma en que se aplica la inteligencia de los datos en la administración agropecuaria. Analizando las *startups* mencionadas en el presente documento en las que el *Big Data* se está aplicando a la gestión agrícola y ganadera en Argentina, **se puede señalar que los puntos más importantes que tiene en común entre todas ellas:**

- Llevan adelante una relación con su segmento objetivo de forma personalizada y de asistencia continua;
- En su gran mayoría, los clientes actuales y potenciales pertenecen al segmento B2B;
- Para desarrollar sus sistemas se nutren de información disponible en fuentes gratuitas, además de complementarla con datos extraídos de la operatoria habitual del cliente;
- Se caracterizan por ofrecer cómo aspecto diferencial del negocio la integración de datos heterogéneos y la integración óptima de información disponible en diversas fuentes;
- El rango de actividades claves que impulsan son la adquisición, análisis, procesamiento y posterior visualización de la información;
- Su modelo de ingresos está basado en la suscripción de clientes a los portales y la posterior venta de activos;
- Para el surgimiento, se financiaron a través de ser ganadoras de premios, subvenciones públicas y privadas y mediante el modelo de *venture capital*;
- Todas ellas se relacionaron para su desarrollo con empresas privadas, centros de investigación, institutos de ciencia y tecnología, además de incubadoras y aceleradoras.

Dadas éstas características que pueden hacer de dicha tecnología un verdadero impacto **se puede extraer los siguientes potenciales beneficios:**

- Mayor producción debido a la identificación de ineficiencias;
- Ahorro de costos y de recursos, gracias al acceso de información personalizada y predictiva sobre las necesidades hídricas, de fertilización, riesgos de plagas y enfermedades, etc., favoreciendo así la preservación de los recursos naturales;
- El productor no sólo tiene a su disposición los datos que obtiene de su actividad, sino que además su base de datos integra información de distintas fuentes, tanto públicas como privadas;
- Visión global y en tiempo real de todos los procesos productivos, sus relaciones y los puntos críticos;
- Servicio accesible *online* y *offline*;
- Toda la información es almacenada en las plataformas, lo cual garantiza la durabilidad de la información y
- Generación de información personalizada a través de indicadores de gestión, comparando operativas y KPIs para diferentes variables.

Los potenciales beneficios del uso de *Big Data* en el sector agropecuario son obviamente considerables, por ejemplo mediante la identificación de ineficiencias en la producción, reducción de costos y de recursos gracias al acceso de información personalizada y predictiva. Pero, simultáneamente se encuentran factores altamente limitantes como la carencia de recursos humanos capacitados y soportes tecnológicos de infraestructura adecuados. Por esto todavía no se ha logrado desde el sector simplificar la tarea a los fundadores de estas nuevas *startups* para poner en práctica variables claves para la expansión de los servicios.

Por su parte, para que los agricultores puedan utilizar estas tecnologías **se deben superar los siguientes retos:**

- Fomentar la inserción de nuevos estudiantes y jóvenes profesionales para que desarrollen capacidades y aptitudes aplicables a las técnicas de análisis y procesamiento de datos masivos;
- Alcanzar beneficios rentables. Es decir, poder cuantificar un *return of investment* positivo que permita la evolución de los proyectos a largo plazo (bajo nivel de riesgo para los productores);
- Reducir los costos de los activos físicos y tiempos de implementación para integrar los productos a la operatoria habitual del cliente;
- Mitigar la competencia de productos importados a través del ofrecimiento de un mayor valor agregado a los servicios no transables, mejorando el soporte local y aumentando la experiencia del usuario;
- Aumentar la conectividad en áreas rurales para facilitar la recepción de los aparatos móviles que dificultan la comunicación entre equipos;

- Falta de entendimiento del productor de las herramientas y sistemas entregados por las empresas, aumentando el tiempo de capacitación y asistencia personalizada para lograr el uso autónomo por parte de los clientes;
- Falta de interpretación de la información brindada. Por ejemplo la aplicación puede indicarle al productor cuánto trigo/soja está produciendo un lote en particular o cuánto y cómo se está alimentando parte de su ganado pero es mucho más complejo entender el por qué se está produciendo de esa manera y que cambios se deben ajustar para la próxima cosecha o cría de ganado.

Entonces se puede concluir, en base a los beneficios, retos y características de las *startups* desarrollados anteriormente que se manifiesta una grieta entre los requerimientos necesarios para desarrollar los proyectos y la situación actual del sector agropecuario argentino, en cuanto a las necesidades técnicas y tecnológicas y por sobre todo en cuanto a la visión del sector en el uso de éstas nuevas tecnologías.

Actualmente se puede recoger una gran cantidad de información, pero la evaluación de la calidad de la misma y la posterior transformación para utilizarla en la toma de decisión es un proceso complejo en el cual se deben desarrollar metodologías y sistemas simples y robustos para que el agricultor pueda obtener toda la capacidad disponible.

El potencial de *Big Data* aplicado al sector agropecuario también está estrechamente relacionado al rol que juegue el agricultor en la demanda de estos servicios para el soporte en la gestión agraria. Además de la predisposición que exista en las instituciones públicas y en las empresas para promover políticas que incentiven el camino del uso intensivo de datos digitales tanto en infraestructura y en formación técnica para el desarrollo de nuevos recursos humanos capacitados para trabajar en el área.

Argentina cuenta con la oportunidad de dinamizar el mercado de *Big Data* y que estas nuevas soluciones se sigan desarrollando para lograr mayor competitividad como país y de ésta forma fortalecer su estructura económica.

Sin embargo, luego de lo expuesto surgen interrogantes que no fueron abarcados en el enfoque del presente trabajo. Por ejemplo saber cómo se están configurando las *startups* en otros sectores de la economía como el financiero, en el sector industrial e incluso para el sector minero y petrolero. En estas áreas ¿compartirán los mismos retos y características que en el sector agropecuario o la dinámica del sector impone diferentes configuraciones a las *startups*? ¿Se obtendrán similares beneficios que los que se obtienen del campo argentino? Incluso también surgen incógnitas sobre la implementación de proyectos de *Big Data* en sectores de la economía que no tengan las mismas estructuras como para sostener este tipo de tecnologías, por ejemplo las

PyMES industriales del Gran Buenos Aires, teniendo en cuenta la importancia que tienen dentro del área tanto con su aporte económico a las arcas de la provincia como en generación de empleo ¿se podrán generar proyectos con costos de implementación menores que sean sostenibles a largo plazo para PyMEs textiles o siderúrgicas? ¿Tendrán la misma reacción los dueños de las fábricas que los propietarios de los campos ante un proyecto de *Big Data*?

6. Anexos

6.1 Encuesta *startup* agro *Big Data* 2018

A. ¿Cuál cree que son los factores limitantes al momento de vender al cliente los servicios?

- Costo de montar y mantener una solución de *Big Data*
- Justificación de un ROI en el negocio para el cliente
- Falta de entendimiento de parte del cliente
- Falta de garantía de satisfacción
- Otros

B. Según su criterio ¿Cuál de los siguientes barreras de entrada cree usted que es la más fuerte para la implantación de empresas vinculadas al *Big Data*?

- RRHH especializados disponibles
- Falta de financiación
- Barreras relacionadas con la capacidad tecnológica
- Falta de criterio sobre la rentabilidad que pueden ofrecer este tipo de inversiones
- Otros

C. De los distintos aspectos diferenciales de su negocio ¿Cuáles considera más importante?

- Apoyar procesos complejos en la toma de decisiones
- Ser capaz de integrar datos de distintas fuentes
- Ser capaz de conjugar datos heterogéneos
- Ser capaz de extraer inteligencia de la combinación de fuentes estructuradas y no estructuradas de información
- Otros

D. ¿Qué tipos de fuente de información utilizan?

- Interna
- Adquirida de otras empresas
- Provisionadas por el cliente
- Disponible gratuitamente
- Otros

E. ¿Qué tipo de financiamiento han utilizado para llevar a cabo la empresa?

- Préstamos participativos
- *Crowdfunding*
- Venture Capital
- Subvenciones publicas
- Subvenciones privadas
- Premios
- Préstamos bancarios
- Programas de apoyo
- Otros

F. ¿Con qué tipos de redes estuvieron vinculadas durante el lanzamiento?

- Incubadoras

- Aceleradoras
 - Instituciones de Ciencia y Tecnología
 - Centros de Investigación
 - Relaciones con empresas privadas
 - Relaciones con empresas publicas
 - Otros
- G. ¿Quiénes son sus clientes (actuales o potenciales)?**
- B2B
 - B2C
 - Ambos
 - Otros
- H. ¿Qué relación tienen con su segmento objetivo?**
- Asistencia personalizada
 - Autoservicio
 - Comunidad
 - Co-creación
 - Otros
- I. ¿Cuál de estas actividades considera que es una actividad clave respecto del uso de datos en organización?**
- Generación de datos
 - Adquisición de datos
 - Procesamiento de datos
 - Análisis de datos
 - Visualización de datos
 - Distribución de datos
 - Otros
- J. ¿Cuál es el modelo de ingresos que utiliza su empresa?**
- Venta de activos
 - Por comisión
 - Leasing
 - Licencias
 - Suscripción
 - Publicidad
 - Otros

7. Bibliografía

- Agroapp.* (s.f.). Recuperado el 31 de Mayo de 2018, de <http://agroapparg.com/>
- Agrozone.* (s.f.). Recuperado el 31 de Mayo de 2018, de <http://www.agrozone.ag/>
- Andes, U. d. (2015). *Definición de Indicadores de Desempeño*. Los Andes.
- Arend, C. (2012). *De que forma el Big Data transforma la protección y el almacenamiento de datos*. Buenos Aires: Consultora IDC.
- Barranco, R. (18 de Junio de 2012). *IBM*. Recuperado el 03 de Abril de 2018, de <https://www.ibm.com/developerworks/ssa/local/im/que-es-big-data/>
- Biobot.* (s.f.). Recuperado el 31 de Mayo de 2018, de <https://www.biobot.com.ar/>
- Brigado, M. (13 de Marzo de 2017). *Nxtp.Labs*. Recuperado el 31 de Mayo de 2018, de <http://agtech.nxtplabs.com/2017/03/13/reportaje-a-jairo-a-trad-cofundador-de-kilimo/>
- Castro Gomez, S. (2018). *Big Data: aprovechando los datos, el nuevo gran activo*. Medellin: ASOBANCARIA.
- Comercio), C. (. (14 de Noviembre de 2017). Ecosistema del emprendedor. *La Nación*, 49.
- CTM Data.* (s.f.). Recuperado el 31 de Mayo de 2018, de <http://ctm-data.com/es/sample-page/#contact>
- Dean, J. y. (2004). *MapReduce: Simplified Data Processing on Large Clusters*. California: Google Inc.
- Derteano, A. (2014). *Big Data en el sector financiero español*. Madrid: EY.
- Empresa-Concepto. (2012). *Creación de empresas con base tecnológica de origen académicas*. Galicia: Adumbro.
- Entrepreneurship, C. d. (2016). *GEM 2015 - Resumen Argentina*. Buenos Aires: IAE.
- Eric, R. (2011). *El método Lean Startup*. Estados Unidos: Crow Publishing Group.
- Farmintech.* (s.f.). Recuperado el 01 de Mayo de 2018, de <https://www.farmintech.com/>
- Farmintech.* (s.f.). Recuperado el 31 de Mayo de 2018, de <https://www.farmintech.com/>
- Fayyad, U. (1996). *Data Mining and Knowledge Discovery in Databases*. Washigton: Microsoft Research .
- Galimany, A. (2014). *La creación de valor en las empresas a través del Big Data*. Barcelona: Universitat de Barcelona.
- Garcia, C. (23 de Agosto de 2016). *Indra*. Recuperado el 2018 de Abril de 21, de <https://www.indracompany.com/es/blogneo/deep-learning-sirve>

- Hernández Leal, E. J. (2016). *Aplicación de técnicas de análisis de datos y administración de Big Data ambientales*. Medellín: Universidad Nacional de Colombia.
- Hernandez, J. (27 de marzo de 2017). *Designplus*. Recuperado el 6 de abril de 2018, de <https://designplus.co/es/blog-marketing-digital/kpis-para-medir-tu-sitio-web>
- Hernandez, M. (09 de Febrero de 2017). *Decideo*. Recuperado el 31 de Mayo de 2018, de https://www.decideo.com/Biobot-Balance-inteligente-para-el-riego-residencial-profesional-y-el-agro_a1649.html
- Hernández-Leal, E. &.-M.-C. (2017). *Big Data: an exploration of research technologies and application cases*. Medellín: TecnoLógicas.
- IBM Institute for Business Value. (2012). *Analytics: el uso del big data en el mundo real*. Madrid: IBM España.
- INTA. (2011). *Evolución del sistema productivo agropecuario argentino*. Buenos Aires: Ministerio de Agricultura, Ganadería y Pesca.
- Josh, J. (8 de junio de 2012). *Domo*. Recuperado el 6 de abril de 2018, de <https://www.domo.com/blog/how-much-data-is-created-every-minute/>
- Kilimo*. (s.f.). Recuperado el 31 de Mayo de 2018, de <http://www.kilimo.com.ar/>
- Liu, M. C. (2014). *Big Data: A survey*. New York: Springer Science Business Media.
- Malvicino, F. (2015). *Descubriendo Big Data en Argentina. Encuesta Digital 2014*. Buenos Aires: AGRANDA.
- Mayer-Schonberger, V. y. (2012). *Big Data: La revolución de los datos masivos*. Madrid: Turner Publicaciones S.L.
- Mejia, J. (1 de abril de 2013). *Juan Carlos Mejia LLano*. Recuperado el 6 de abril de 2018, de <http://www.juancmejia.com/marketing-en-redes-sociales/social-big-data-definicion-e-importancia-de-big-data-en-redes-sociales/>
- Pascual Fuste, N. (2017). *Estudio de empresas aceleradoras e incubadoras españolas*. Valencia: Universidad Politécnica de Valencia .
- Pisani, N., & Miazzo, D. (2017). *El campo argentino en números*. Rio Cuarto: FADA.
- Prado, A. I. (2012). *Cómo la avalancha de datos se ha convertido en un importante beneficio*. Madrid: TicBeat.
- Regunaga, M., & Tejeda, A. (2015). *La política de comercio agrícola de Argentina y el desarrollo sustentable*. Suiza: ICTSD.
- Sagiroglu, S. &. (2013). *Big Data: A Review*. Ankara: Faculty of Engineering Turkey.
- Sampieri Hernandez, R. (1997). *Metodología de la investigación*. Mexico: McGraw Hill Interamericana de Mexico.

- Serrano, E. (12 de Enero de 2017). *Agro Inteligencia*. Recuperado el 15 de Abril de 2018, de <http://www.agrointeligencia.com/big-data-agricultura-enrique-serrano/>
- Simapro*. (s.f.). Recuperado el 31 de Mayo de 2018, de <https://www.monitoreoagricola.com/prices>
- Simarro, D. (12 de Abril de 2017). *Ainia*. Recuperado el 15 de Abril de 2018, de <http://www.ainia.es/tecnoalimentalia/tecnologia/big-data-y-agricultura-de-precision-como-reducir-los-riesgos-en-la-produccion-de-alimentos-agropecuarios/>
- Sismagro*. (s.f.). Recuperado el 31 de Mayo de 2018, de <https://www.sismagro.com/funcionalidades-de-sismagro/>
- Wong Cheang, J. C. (2005). *Ley de Moore, Nanotecnologías y Nanociencias*. Mexico: Revista Digital Universitaria.
- Yuniet, S. (2009). *Herramientas de Minería de Datos*. Habana: Revista Cubana de Ciencias Informáticas.