

PROBLEMATIZACIÓN DE LA APLICACIÓN DEL BIG DATA EN EL CASO SISTEMA UNIFICADO DE INFORMACIÓN OBLIGATORIA DE LAS OPERACIONES DE COMPRAVENTA DE GRANOS (SIO-GRANOS)¹

Esteban Barelli

Asociación Argentina de Consorcios Regionales de Experimentación Agrícola (AACREA). Sarmiento 1236, C1012 CABA. República Argentina.

ebarelli@crea.org.ar

Resumen

<p>Recibido: 12/2018</p> <p>Aceptado: 04/2019</p>	<p>El “<i>Big data</i>” es una realidad cotidiana en la vida de las personas y organizaciones del sector público y privado. Provista de numerosas definiciones y alcances, la potencialidad del “<i>Big data</i>” en brindar soluciones y servicios también conlleva un riesgo por las implicancias que podría tener en la sociedad.</p> <p>El sector agroindustrial de Argentina cuenta hoy con el gran desafío de adoptar, ajustar y aplicar procesos y tecnologías de “<i>Big data</i>” que sean capaces de generar soluciones y oportunidades, aportando a la toma de decisiones con un real impacto en los resultados. Por tal motivo, es importante relevar las diferentes fuentes y herramientas que puedan brindar datos para conformar una efectiva aplicación de “<i>Big data</i>”, analizando no solo las potencialidades sino también sus riesgos e implicancias para con los propios productores, los demás actores de la cadena y el resto de la sociedad y organizaciones (públicas y privadas).</p> <p>En este trabajo se presenta la problematización de la aplicación de “<i>Big data</i>” en un caso del sector agroindustrial, específicamente en el mercado de granos a través de la metodología de las “4 R” en la iniciativa del SIO-Granos.</p>
<p>Palabras clave</p> <p>Big data.</p> <p>Externalidad.</p> <p>Agroindustrial.</p> <p>SIO granos.</p>	

Copyright: Facultad de Ciencias Económicas, Universidad de Buenos Aires.

ISSN: 2250-687X - ISSN (En línea): 2250-6861

¹ Los conceptos y opiniones contenidos en este trabajo son de exclusiva responsabilidad de la autora.

PROBLEMATIZATION OF THE IMPLEMENTATION OF BIG DATA IN THE CASE OF THE MANDATORY INFORMATION ON GRAIN TRADING OPERATIONS UNIFIED SYSTEM (SIO - GRAIN)

Esteban Barelli

Asociación Argentina de Consorcios Regionales de Experimentación Agrícola (AACREA). Sarmiento 1236, C1012 CABA. República Argentina.

ebarelli@crea.org.ar

Abstract

KEYWORDS	Big data is a reality in both people and organization in public and private sector. Big data concept has many definitions and scope. Its potentiality to provide solutions and services also implies risks owing to its implications over the society.
Big data.	The agroindustrial sector of Argentina faces a big challenge in order to adopt, adjust and apply Big data's process and techs able to generate solutions and opportunities, supporting the decision making process with a real impact on the results. For that reason it is important to search the different sources and tools that can provide data in an effective Big data process, analyzing not only the potentiality but the risks and implications for the producers and the rest of the actors (value chain), the society and organizations (public and private). This paper presents the approach of the Big data's potential problems in an agroindustrial sector case, focusing on the grain market (" <i>SIO granos</i> ") through the "4 R" methodology.
Agroindustrial.	
Grain market.	

Copyright: Facultad de Ciencias Económicas, Universidad de Buenos Aires.

ISSN: 2250-687X - ISSN (En línea): 2250-6861

INTRODUCCIÓN

El “*Big data*” es una realidad cotidiana en la vida de las personas y organizaciones del sector público y privado. Provista de numerosas definiciones y alcances, la potencialidad del “*Big data*” en brindar soluciones y servicios también conlleva un riesgo por las implicancias que podría tener en la sociedad. Dichas implicancias no están claramente identificadas sino que tienen que ver con un contexto general y aplicaciones particulares. Así, considerando el potencial que presenta el “*Big data*” para generar oportunidades, es menester contar con herramientas que permitan caracterizar ese contexto para tener una aproximación a los riesgos que implica aplicar las técnicas y procesos del “*Big data*”.

El sector agroindustrial de Argentina cuenta hoy con el gran desafío de adoptar, ajustar y aplicar procesos y tecnologías de “*Big data*” que sean capaces de generar soluciones y oportunidades, aportando a la toma de decisiones con un real impacto en los resultados. Por tal motivo, es importante relevar las diferentes fuentes y herramientas que puedan brindar datos para conformar una efectiva aplicación de “*Big data*”, analizando no solo las potencialidades sino también sus riesgos e implicancias para con los propios productores, los demás actores de la cadena y el resto de la sociedad y organizaciones (públicas y privadas).

En este trabajo se presenta la problematización de la aplicación de “*Big data*” en un caso del sector agroindustrial, específicamente en el mercado de granos a través de la iniciativa del SIO-Granos. El SIO-Granos es una acción público privada que busca transparentar la comercialización de granos, siendo una plataforma de datos e información capaz de contribuir a caracterizar la configuración del mercado de granos argentino. De este modo, para problematizar el caso SIO-Granos en primer lugar se relevar algunas definiciones referidas al “*Big data*”, incluyendo las potencialidades y riesgos. En ese sentido, se plantean los riesgos desde el concepto de externalidad, se propone un método para contextualizar los mencionados riesgos y sus implicancias (método 4R). Por último se presenta un breve repaso del contexto normativo argentino respecto a la gestión de datos personales, para después abordar la aplicación del método 4R en el caso del sector agroindustrial elegido: SIO-Granos.

1. DEFINICIONES

Si bien en la actualidad el “*Big data*” está instalado en los diferentes ámbitos de la vida cotidiana, es oportuno en esta instancia hacer un repaso de sus conceptos y alcances para luego así poder abordar el caso de estudio planteado en este texto.

Una primera definición de “*Big data*” la podríamos encontrar en lo propuesto por Boyd y Crawford (2012). Ellos hacen hincapié en la capacidad de investigar, agregar y relevar una gran cantidad de datos. Además indican que es un fenómeno cultural, tecnológico y académico que tiene su sustento en tres intersecciones: tecnológico (gestión de datos), análisis (búsqueda e identificación de patrones) y mitológico (la creencia de la potencialidad del “*Big data*”). Kitchin (2013) lo caracteriza de la siguiente manera:

- Gran volumen de datos (“*tera o petabyte*”);
- Alta velocidad de generación (actualización);
- Diversidad en variedad;
- Alcance general (“*n=all*”);

- Precisión en los datos;
- Relación con otros conjuntos de datos;
- Flexible y escalable.

En tanto “*Data mining*” se podría definir operacionalmente como el proceso para descubrir patrones estructurales, en forma automática o semiautomática dentro de un gran volumen de datos, y donde dichos patrones son útiles resolviendo un problema (Witten, Frank, Hall, & Pal, 2016). Es adquirir una descripción estructural (descubrir patrones) de la muestra de datos (ejemplos) que se está describiendo o analizando (McQueen, Garner, Nevill-Manning, & Witten, 1995). Del mismo modo, Witten et al. (2016) definen “*Machine learning*” como la manera en que se lleva adelante el “*Data mining*” y cómo éste aprende cambiando su comportamiento en una forma tal que le permite tener un comportamiento mejor en el futuro. Es decir, el “*Machine learning*” tiene que ver más con la performance que con el stock de conocimiento. Implica pensar, implica propósito o intencionalidad (a diferencia del “*training*”).

Planteadas las definiciones, se desprende la gran potencialidad del “*Big data*” para poder utilizar sus capacidades en resolver problemas y brindar nuevos productos y servicios, tanto a niveles micro como niveles agregados y en el ámbito público como privado. Así, Schmarzo (2013) establece el desafío de ser capaz de utilizar y apalancarse en los datos e información de clientes y/o productos originados en el “*Big Data*” para generar valor en las empresas y organizaciones. Según Steinmann, Matei, y Collmann (2016), dichas capacidades (metodologías) del “*Big data*” tienen un doble potencial. En primer lugar, se podría identificar en la aplicabilidad descripta anteriormente (aportar soluciones y generar valor), y una segunda potencialidad es debido a la generación de dilemas éticos originados en los procesos y resultados obtenidos al aplicar las capacidades y alcances del “*Big data*”. Además, en esta última potencialidad deberíamos tener en cuenta otras externalidades negativas que se podrían derivar de la generación y procesamientos de dichos datos e información por terceros.

En conjunto, dilemas éticos y externalidades negativas (en adelante externalidades) son elementos de relevancia y se deberían considerar al momento de plantear una iniciativa “*Big data*”, cualquiera sea su disciplina y ámbito de aplicación. Así, a continuación haremos un repaso general de las principales externalidades relevadas, para luego aplicar un abordaje metodológico al caso SIO-Granos teniendo en cuenta el actual contexto legal argentino.

2. EXTERNALIDAD

Tomando la sugerencia de Cuquet, Vega-Gorgojo, Lammerant, y Finn (2017), a los fines del presente trabajo, externalidad es definida de acuerdo a lo establecido por la Organización para la Cooperación y el Desarrollo Económicos (OCDE): “*situación donde el efecto de la producción o consumo de un bien y servicio implica costos o beneficios en otros que no están reflejados en el precio cargado en el bien o servicio que ha sido provisto*” (Transition, 1993). Ampliándose las implicancias en el ámbito económico, social, ético y político, y no solo circunscripto en el presente, sino también aquellos potenciales riesgos en el futuro. En ese sentido, es oportuno mencionar a Nissenbaum (2011). El autor pone énfasis en el contexto del uso del “*Big data*”, y plantea que el contexto es el que determinará qué tipo de análisis y principios éticos conlleva, es decir las potenciales externalidades. De esta manera, para tener una aproximación a las externalidades de la aplicación de “*Big data*” y así poder contextualizar su potencial problemática en el caso SIO-Granos se propone la metodología planteado por Steinmann et al. (2016).

2.1 El método 4R

Según los autores, la externalidad del “*Big data*” se fundamenta principalmente en dos preocupaciones: la tendencia a la exactitud y precisión de los datos (potencialidad de conocer comportamiento humano -y no humano también-) y la posibilidad de las 4R “*reuse, repurpose, recombine, or reanalyze*”, lo que conlleva a otros usos y amenazas para la población/usuarios ya que se producen nuevos conocimientos/usuarios. A los fines de este trabajo, se supone que la primera preocupación (exactitud y precisión) se encuentra contenida de alguna forma en la segunda preocupación (4R), por lo que, a continuación se desarrolla el marco de las 4R que será la propuesta de abordaje metodológico para analizar el caso SIO-Granos.

Para la primer R (“*reuse*”), se refiere a la recolección de datos e información que originalmente se utilizaron para un estudio en particular y luego se utiliza nuevamente con fines y en un campo de investigación comparables al original. El dilema está en la responsabilidad que tienen los científicos que en primer lugar la utilizaron (en EE.UU. bajo la “*Common Rule*”² –“*Office for Human Research Protection 1993*”-) sobre los usos por parte de terceros.

Diferente es el caso cuando se utilizan los datos e información con otros fines (segunda R, “*repurpose*”). Con la misma interrogante respecto a la responsabilidad de los terceros en la nueva utilización de los datos, en esta instancia se suma la legitimidad de analizar dichos datos que fueron solicitados/requeridos bajo un contexto de privacidad totalmente diferente. Hay algunas iniciativas en EE.UU. por parte de la “*Common Rule*” a través del “*Notice of Proposed Rulemaking (NPRM)*” pero todavía no hay definiciones³.

La recombinación de datos e información implica la utilización de diferentes fuentes primarias que, en el contexto de “*Big data*”, pueden ser relacionadas y recombinadas generando nueva información y conocimientos, capaces de cambiar el contexto de un mercado. El riesgo es la potencialidad de identificar o “*trackear*” personas, empresas y organizaciones, incorporando un nuevo interrogante a la amenaza de la privacidad de las personas.

Por último, el “*Big data*” supone la creación de grandes bases de datos que son utilizadas bajo un objetivo específico que surge del estudio llevado a cabo. Dichas bases de datos en un futuro pueden estar disponibles para nuevos análisis a medida que se generen nuevos conocimientos en la materia. El cruzamiento de bases de datos podría causar acciones deliberadas (discriminación) contra personas, empresas y organizaciones. Así, en esta instancia es clave la concientización de los proveedores de datos que su información puede ser utilizada con diferentes fines al original.

Si bien los alcances se plantean en cuatro vías (“*reuse, repurpose, recombine, or reanalyze*”), no hay una distinción marcada y algunas externalidades pueden estar asociadas a más de una de las vías.

3. CONTEXTO LEGAL ARGENTINO

Para realizar la problematización de la aplicación del “*Big data*”, es menester tener una aproximación al contexto legal argentino en cuanto a normativa aplicable a las externalidades que surgen de su capacidad de procesamiento actual y las nuevas tecnologías. Así, en el plano nacional y referidas al análisis de las leyes y regulaciones específicas existentes orientadas a la protección de datos personales se puede mencionar el artículo 43 de la Carta Magna argentina⁴, como así también

² *Office for Human Research Protections - U.S. Department of Health & Human Services*: <https://www.hhs.gov/ohrp/>

³ *U.S. Department of Health & Human Services*, Comunicación Oficial: <https://www.hhs.gov/ohrp/announcement-nprm-revisions-common-rule.html>

⁴ <http://servicios.infoleg.gob.ar/infolegInternet/anexos/0-4999/804/norma.htm>

la Ley 25.326 de Protección de los Datos Personales (LPDP)⁵ y sus decretos reglamentarios (Riva, Abbateamarco, & Cervino, 2017).

Respecto a la Constitución Nacional, de acuerdo a Scanavino (2012), es el artículo 43 de la Carta Magna que consagra constitucionalmente el “*Habeas data*” ante violaciones a los derechos de privacidad o intimidad de las personas. Así, los particulares afectados pueden reclamar: la corrección de datos inexactos, la eliminación de datos inexactos, superfluos o vinculados a cuestiones personalísimas e íntimas vedadas por la ley 23.592⁶ (antidiscriminatoria), la no divulgación de los mismos, y/o su actualización (Loreti, 1995). Además, es de destacar que el ejercicio de este derecho (privacidad o intimidad) aparece también en pactos internacionales a los que la Argentina ha suscrito y que tienen en nuestro país fuerza constitucional, ya que son consagrados en la Constitución Nacional a través del artículo 75 inciso 22 de la misma (Riva et al., 2017).

Más específicamente en la materia, en el año 2000 se sanciona la Ley Nacional 25.326 LPDP que tiene por objeto “*la protección integral de los datos personales asentados en archivos, registros, bancos de datos, u otros medios técnicos de tratamiento de datos, sean éstos públicos, o privados destinados a dar informes, para garantizar el derecho al honor y a la intimidad de las personas, así como también el acceso a la información que sobre las mismas se registre, de conformidad a lo establecido en el artículo 43, párrafo tercero de la Constitución Nacional*”. Con la reglamentación de la mencionada ley se crea la Dirección Nacional de Protección de Datos Personales (DNPDP)⁷, principal órgano de gestión. A los fines del presente trabajo, se presentan las definiciones del artículo 2 de mencionada Ley:

— Datos personales: Información de cualquier tipo referida a personas físicas o de existencia ideal determinadas o determinables.

— Datos sensibles: Datos personales que revelan origen racial y étnico, opiniones políticas, convicciones religiosas, filosóficas o morales, afiliación sindical e información referente a la salud o a la vida sexual.

— Archivo, registro, base o banco de datos: Indistintamente, designan al conjunto organizado de datos personales que sean objeto de tratamiento o procesamiento, electrónico o no, cualquiera que fuere la modalidad de su formación, almacenamiento, organización o acceso.

— Tratamiento de datos: Operaciones y procedimientos sistemáticos, electrónicos o no, que permitan la recolección, conservación, ordenación, almacenamiento, modificación, relacionamiento, evaluación, bloqueo, destrucción, y en general el procesamiento de datos personales, así como también su cesión a terceros a través de comunicaciones, consultas, interconexiones o transferencias.

— Responsable de archivo, registro, base o banco de datos: Persona física o de existencia ideal pública o privada, que es titular de un archivo, registro, base o banco de datos.

— Datos informatizados: Los datos personales sometidos al tratamiento o procesamiento electrónico o automatizado.

⁵ <http://servicios.infoleg.gob.ar/infolegInternet/anexos/60000-64999/64790/norma.htm>

⁶ <http://servicios.infoleg.gob.ar/infolegInternet/anexos/20000-24999/20465/texact.htm>

⁷ <https://www.argentina.gob.ar/aaip/datospersonales>

— Titular de los datos: Toda persona física o persona de existencia ideal con domicilio legal o delegaciones o sucursales en el país, cuyos datos sean objeto del tratamiento al que se refiere la presente ley.

— Usuario de datos: Toda persona, pública o privada que realice a su arbitrio el tratamiento de datos, ya sea en archivos, registros o bancos de datos propios o a través de conexión con los mismos.

— Disociación de datos: Todo tratamiento de datos personales de manera que la información obtenida no pueda asociarse a persona determinada o determinable.

Continuando en la Ley, el artículo 3 se centra en la licitud de las bases de datos, debiendo estar inscriptos, y tener fines no contrarios a ley y la moral. En tanto el artículo 4 es relevante en cuanto a la calidad de los datos: veracidad y de acuerdo al objetivo que se busca, recolección leal, utilizados para un fin específico, actualizados de ser necesarios, los datos inexactos suprimidos y/o sustituidos, un almacenamiento que permita acceso de su titular y deben ser destruidos una vez finalizado su fin. El artículo 5 es clave dado que hace foco en el consentimiento libre, expreso e informado del titular de los datos. Presentan excepciones aquellos casos por acceso público irrestricto, debido a que se recaben por el ejercicio de las funciones propias del Estado, sean operaciones comerciales enmarcadas en el artículo 39 de la Ley 21.526⁸, se deriven de una relación contractual y/o se limiten a nombre, documento nacional de identidad y datos previsionales. La información provista a los dueños de los datos al momento de la solicitud se establece en el artículo 6: finalidad (y sus destinatarios finales), existencia del archivo (y su dirección y responsable), consecuencias de brindar los datos y la posibilidad de acceso, rectificación y/o supresión de los datos. Estos últimos tienen sanciones en caso de no cumplimiento de acuerdo a los artículos 31 (administrativas) y 32 (penales) de la Ley. En tanto el artículo 9 hace mención a la seguridad de los datos, donde los responsables deben adoptar las medidas necesarias para el resguardo de la confidencialidad y la seguridad de los datos. Si bien el decreto original (1558/2001) no dice mucho, reglamentaciones posteriores⁹ han establecido niveles de seguridad, medidas y modelos para el registro (Riva et al., 2017). Por último, una excepción se menciona en el artículo 28: *“Las normas de la presente ley no se aplicarán a las encuestas de opinión, mediciones y estadísticas relevadas conforme a Ley 17.622¹⁰, trabajos de prospección de mercados, investigaciones científicas o médicas y actividades análogas, en la medida que los datos recogidos no puedan atribuirse a una persona determinada o determinable”*.

Además del marco regulatorio legal, hay parámetros o iniciativas que buscan dar respuestas alternativas desde el sector privado ante las externalidades del “Big data”. De esta manera, se van planteando antecedentes a considerar en casos donde la ley no es clara o hay diferentes interpretaciones. Una opción disponible en cuanto a estándares deseables o buscados de seguridad de los datos o sistemas de información es la aplicación de las normas de la familia ISO 27000. Para la protección de los datos y los sistemas de información, los estándares ISO 27000, 27001 y 27002 proveen controles de objetivos específicos, requisitos y normas, con los cuales las organizaciones pueden lograr alcanzar una adecuada seguridad (Disterer, 2013).

⁸ <http://servicios.infoleg.gob.ar/infolegInternet/anexos/15000-19999/16071/texact.htm>

⁹ Resoluciones 11/2006 y 09/2008 que aprueban las Medidas de Seguridad para el Tratamiento y Conservación de los Datos Personales Contenidos en Archivos, Registros, Bancos y Bases de Datos Públicos no estatales y Privados.

¹⁰ <http://servicios.infoleg.gob.ar/infolegInternet/anexos/20000-24999/24962/texact.htm>

4. CASO SIO-GRANOS

Aún con matices y particularidades propias de cada actividad que lo integra, el sector agroindustrial está atravesando una revolución digital que es interesante para hacer foco por su alcance e impacto. Un ejemplo, incluso productores de baja escala reúnen información que es recolectada por equipos de agricultura de precisión, y muchos productores están usando información y análisis de grandes bases de datos para la toma de decisiones (Bronson & Knezevic, 2016). En Argentina, entre dichas bases de datos se destaca por su potencialidad la iniciativa público privada SIO-Granos.

El Sistema unificado de Información Obligatoria de las operaciones de compraventa de granos (SIO-Granos) que conforman el Mercado Físico surge de acuerdo a lo establecido en la Resolución Conjunta CNV-SAGyP N° 208 y 628 ¹¹. El mencionado sistema ha sido desarrollado por diversas instituciones involucradas en la comercialización agrícola ¹² y su autoridad de aplicación es la Secretaría de Agroindustria (SecAgro) a través de su Secretaría de Mercados Agroindustriales. Su principal objetivo es mejorar la transparencia en el mercado de granos.

De acuerdo a la información disponible ¹³, el SIO-Granos permite reflejar diariamente las diferentes operaciones concertadas de compra venta de granos ¹⁴ a nivel país. Los datos relevados presentan el siguiente detalle: fecha de concertación, operación, tipo, cantidad, producto (grano), calidad, zona geográfica de procedencia y entrega, precio, fecha de entrega y condición de pago. Así, dichas operaciones son declaradas por el comprador y/o intermediario autorizado ¹⁵ y cargadas al sistema del SIO-Granos (Scanavino, 2012). Una vez cargadas al sistema, son procesadas y ordenadas de acuerdo a las características detalladas, donde luego compradores y demás actores de la comercialización granaria pueden consultar a través de un monitor ¹⁶ los precios y condiciones vigentes. De esta manera, el SIO-Granos se presenta como una oportunidad interesante para relevar lo que Garcia-Fronti (2013) denomina “configuración del mercado”: *“la existencia de una multiplicidad de formas prácticas de confrontación entre la oferta y la demanda para definir un mercado”*. Estas configuraciones pueden tener una gran relevancia a la hora de definir precios y relaciones comerciales (Garcia-Fronti, 2013).

Así, contar con los detalles de las operaciones de compra venta de granos del país (características claves de la configuración de un mercado), y en conjunto con otros datos, permitiría encontrar patrones de comportamiento en la comercialización por parte de los productores y la demanda, sin buscar en ningún momento la identificación de los actores. Dichos patrones podrían resultar de interés al momento de la planificación y tomar decisiones por parte de los propios productores (riesgo de mercado y logística), de empresas que brindan productos/servicios al sector (oportunidades de ventas) y el sector público (asignación de recursos y propuestas de política pública). Así, dada la potencialidad de la aplicación del “Big data” en esta fuente de datos, a continuación se presenta la problematización de avanzar en dicho planteo respecto a las posibles externalidades negativas en el caso SIO-Granos.

¹¹ <http://servicios.infoleg.gob.ar/infolegInternet/anexos/230000-234999/232154/texact.htm>

¹² Las Bolsas de Cereales de Bahía Blanca, Buenos Aires, Córdoba y Entre Ríos y las Bolsas de Comercio de Chaco, Rosario y Santa Fe, el Mercado a Término de Buenos Aires S.A. (MATBA) y el Mercado a Término de Rosario S.A. (ROFEX).

¹³ Web institucional: <https://www.siogranos.com.ar/index.html#2>

¹⁴ Granos relevados: cebada (forrajera), girasol, maíz, trigo (pan y candeal), soja (y aceite) y sorgo.

¹⁵ Inscriptos en el Registro Único de la Cadena Agroalimentaria (RUCA).

¹⁶ Monitor SIO-Granos: <http://monitorsiogranos.magyp.gob.ar/monitorsiogranos.html>

Caso SIO-Granos, potenciales externalidades

En primer lugar, y de acuerdo a las definiciones de la Ley 25.326 LPDP, el planteo general del caso SIO-Granos implicaría la recolección de datos de operaciones comerciales de granos realizadas por personas físicas y jurídicas (sin identificación directa), la creación de un nuevo archivo y la realización de un tratamiento de los datos recolectados que permita de este modo encontrar los patrones de comportamiento del mercado granario. Si bien no hay detalle directo alguno que permita la identificación de vendedores/compradores, se plantearía la práctica de disociación de datos en cualquier instancia que haya peligro de externalidad y que pueda afectar de este modo privacidad o permita identificar personas físicas/jurídicas. Además de la recopilación y tratamiento de los datos aportados por el SIO-Granos, también se espera recopilar y procesar otros datos de fuentes distintas que aporten la evolución de variables claves que sean capaces de ayudar a detectar patrones en el mercado granario. Ninguna de ellas es de carácter personal, sino que refieren a variables de índole económica y/o productiva (precios de insumos, productos y servicios o precipitaciones entre otras). Por último, no implicaría bajo ningún concepto la recopilación, archivo y tratamiento de datos sensibles correspondientes a ninguno de los actores involucrados en la operatoria diaria del SIO-Granos.

Habiendo dicho el planteo general y sus principales implicancias, para analizar las potenciales externalidades y así abordar la problemática, se contextualiza el caso SIO-Granos con el método de las 4R “*reuse, repurpose, recombine y reanalyze*”, teniendo de esta manera una aproximación a las potenciales externalidades de aplicar “*Big data*”.

Para la reutilización (“*reuse*”), se debe plantear la problematización desde dos puntos de vista: desde los actuales responsables de la base de datos del SIO-Granos y del futuro responsable de la base de datos generada si se aplica la iniciativa de “*Big data*”.

En el primer caso, dada la definición de reutilización, datos que se utilizan nuevamente con fines y en un campo de investigación comparables al original, resulta difícil determinar si hay responsabilidad alguna por parte de los originadores del SIO-Granos. En primer lugar, resulta difícil establecer si la temática del caso “*Big data*” en SIO-Granos no puede ser considerada una investigación con fines y en un campo comparables al original. Si bien en una primera lectura se podría decir que los fines son distintos, se podría argumentar que la investigación para detectar esa “configuración del mercado” podría aportar a la transparencia del mercado, dado que se aportaría mayor información respecto a un producto específico. Conforme mayor sea la disponibilidad de información, mayor será la transparencia que se tenga sobre ese producto (Lam, 2010). De este modo, si sorteáramos la mencionada dificultad y asumimos que los fines son similares, y siguiendo el método 4R, se podría asignar una potencial externalidad de “*reuse*” para los responsables actuales de las bases de datos del SIO-Granos. Responsabilidad que hoy no tendría resguardo en la legislación argentina, tal lo relevado en instancias anteriores del presente trabajo.

En el segundo caso, y al igual que en el anterior, si se aplica “*Big data*” al SIO-Granos y los resultados se toman para estudios con fines y temáticas similares, y así de alguna manera se vulneran los derechos de privacidad y/o se identifican a personas físicas/jurídicas, hoy la legislación argentina no resguardaría a los responsables. De este modo, aunque se podría plantear estrategias de disociación de datos, nuevamente la iniciativa “*Big data*” en SIO-Granos podría implicar una potencial externalidad para aquellos que son los originadores de las bases de datos y los resultados.

Para el caso de “*repurpose*”, la ley vigente podría ser una limitante para la aplicación de “*Big data*” en SIO-Granos debido a que los actores involucrados en la Resolución Conjunta CNV-SAGyP N° 208 y 628, y que generan el relevamiento de precios y condiciones, no habrían dado el consentimiento específico para la utilización de sus datos en la búsqueda de patrones de comportamiento en el mercado granario (objetivo de la aplicación de “*Big data*” en SIO-Granos). Esta cuestión es central y se establece claramente en la Ley 25.326 que dispuso dos requisitos básicos que son condición de licitud de todo tratamiento de datos: a) requerir el consentimiento previo del titular del dato (artículo 5°), y b) brindar información al titular del dato (artículo 6°).

Sin embargo, y sumado a la dificultad mencionada en el anterior ítem con respecto a la argumentación de considerar la búsqueda de patrones como acción que podría también tener como objetivo transparentar el mercado (al igual que SIO-Granos), la norma vigente (Resolución Conjunta CNV-SAGyP N° 208 y 628) tampoco establece restricciones al uso de los datos publicados, siendo enfáticos en la difusión y uso: “*Que la difusión masiva de las operaciones y el acceso público a información completa, sienta las bases para una equitativa competitividad a ser aplicada en la concertación de los contratos bajo las condiciones más favorables para las partes, con especial atención en el fortalecimiento de la protección del pequeño productor, industrial e inversor*”. De este modo, la externalidad por la utilización de los datos con otros propósitos no es tajante, tanto si se considera o no a la iniciativa de “*Big data*” en SIO-Granos como una acción que también busca transparentar mercados (al igual que el SIO-Granos) como si se enmarca la aplicación del “*Big data*” como una iniciativa que pretende promover y difundir el SIO-Granos tal lo expresa y pretende la Resolución Conjunta. Así, de acuerdo a cual posición sea considerada, el potencial riesgo será distinto e incluso podría estar exento de externalidad.

Por último, en esta vía además se suma la potencial utilización de terceros con los resultados obtenidos por el caso de aplicar “*Big data*” al SIO-Granos, teniendo que considerar otra vez de esta forma la potencial responsabilidad por dicha utilización, hoy sin resguardo por la ley argentina.

En recombinación (nuestra tercer R), tienen particular importancia en el presente caso de aplicación de “*Big data*” y por el principio de finalidad, lo que la Ley 25.326 dispone en su artículo 4: “*los datos objeto de tratamiento no pueden ser utilizados para finalidades distintas o incompatibles con aquellas que motivaron su obtención*” (Brussa, Castillo, Ron, & Gerke, 2017). No solo es relevante lo establecido en dicho artículo (finalidad, que podría ser discutible de acuerdo a lo planteado en las erres anteriores) sino también principalmente los artículos 5 (consentimiento) y 6 (información al dueño de los datos). De este modo, si se aplica “*Big data*” al caso SIO-Granos no se estaría cumpliendo ninguno de los artículos anteriormente mencionados ya que los dueños de los datos no habrían dado un consentimiento libre, expreso e informado, y además no se habría brindado la información previamente a los titulares de forma expresa y clara. Así, en una primera aproximación en la recombinación hay un claro potencial de externalidad a tener en cuenta.

Otra consideración en la recombinación es respecto a la posibilidad de identificar a personas. Si bien en el SIO-Granos no hay datos que permitan identificar directamente a los actores involucrados, en la recombinación habría un potencial riesgo. Los datos provistos en el SIO-Granos podrían ser cruzados con otras fuentes, generando así nueva información y una posible externalidad en particular por la capacidad de poder identificar tanto los compradores como los vendedores.

El desarrollo y accesibilidad de nuevas tecnologías, tales como la utilización de sensores remotos para clasificar y relevar cultivos o el relevamiento sistemático de condiciones ambientales por citar algunos, hacen este tipo de externalidad como una de las más importantes e impredecibles, tanto para actores del sector público como privado. Es interesante lo dicho por Bronson y Knezevic (2016). Los autores indican que la existencia actual de aplicaciones “*Big data*” en el sector

agroindustrial ha revelado numerosos análisis y herramientas que podrían haber tenido implicancias en la relación de poder entre los actores en el caso del sistema norteamericano. En ese sentido, Boyd y Crawford (2012) también indagan sobre el riesgo de la aparición de una brecha digital entre personas y organizaciones que utilizan o no utilizan la tecnología y procesos de “*Big data*”.

La última R es “*reanalyze*”. En cuanto a la gestión de las bases de datos, el artículo 3 de la Ley 25.326 se ocupa de los requisitos que deberían contemplar para la licitud de archivar datos. Cumpliendo dicha norma para estar habilitado a la tenencia de una base personal de datos, queda latente la externalidad por la posterior utilización pasado un tiempo ante un nuevo contexto científico, tecnológico legal y social.

La generación y disponibilidad creciente de bases de datos de acceso público (y privadas) es una situación que podría originar una externalidad. El desarrollo de nuevas fuentes de datos, adelantos en el procesamiento y la aparición de nuevos usos, podrían presionar para plantear nuevos análisis y procedimientos. Así, bases de datos son reanalizadas generando nuevos contenidos, algunas de los cuales pueden afectar a las personas u organizaciones. Por ejemplo, hay un riesgo que gobiernos usen las distintas fuentes de datos (redes sociales o la plataforma SIO-Granos por ejemplo) y prácticas de “*Big data*” como una forma actual de vigilancia, donde dichas plataformas son simultáneamente canales y fuentes de bases de datos para que ello ocurra (Kennedy & Moss, 2015).

Toma mayor relevancia tanto por el estado actual y en caso de desarrollo del caso SIO-Granos, las posibles implicancias que podría tener por el indebido uso, principalmente por parte del sector público a través de sus diferentes entes y niveles de gobierno, ya que cuentan con diferentes y variadas capas de datos e información de los ciudadanos. En este punto es clave la generación de confianza entre los actores involucrados y los que aplican el “*Big data*”, confianza que en el caso de ser el gobierno (en sus diferentes vías) es de compleja realización por el tipo de vínculo y la propia dinámica del ámbito político.

CONCLUSIONES

El “*Big data*” ya es una realidad en la vida de las personas y organizaciones del sector público y privado. Provista de numerosas definiciones y alcances, la potencialidad del “*Big data*” en brindar soluciones y servicios también conlleva externalidades por sus implicancias que podría tener en la sociedad.

El sector agroindustrial de Argentina cuenta hoy con el gran desafío de adoptar, ajustar y aplicar procesos y tecnologías de “*Big data*” que sean capaces de generar soluciones y oportunidades, aportando a la toma de decisiones con un real impacto en los resultados. Por tal motivo, en este trabajo en primer lugar fue importante tener en claro algunas definiciones, relevar el marco legal vigente en el país, para luego sí profundizar en una potencial fuente de datos de aplicación de “*Big data*”, el caso SIO-Granos. En el mismo se analizó no solo las potencialidades sino su problematización a través de sus riesgos e implicancias para con los propios productores, los demás actores de la cadena y el resto de la sociedad y organizaciones (públicas y privadas).

Así, el contexto legal argentino, y referido al análisis de las leyes y regulaciones específicas existentes orientadas a la protección de datos personales, presenta antecedentes que dan un marco legal que legisla sobre la materia. En ese sentido, se destaca el artículo 43 de la Carta Magna argentina, la Ley 25.326 de Protección de los Datos Personales (LPDP) y sus distintos decretos reglamentarios. Incluso también se podrían mencionar iniciativas privadas que buscan dar respuestas ante las externalidades del “*Big data*”, tales la aplicación de las normas ISO 27000, 27001 y 27002.

Para el caso de la problematización de la aplicación de “*Big data*” en el caso SIO-Granos, se analizó a través de la metodología de las preocupaciones de las 4R desarrollada por Steinmann et al. (2016), siendo las preocupaciones la reutilización, el repropósito, la recombinación y el reanálisis. La aplicación de la metodología permitió identificar las distintas externalidades, e indagar sobre las implicancias que podrían tener para con los responsables de las bases de datos y los dueños de los datos, como así también si el marco legal argentino da respuesta a los desafíos planteados. Se destaca que el planteo general del SIO-Granos supone una recolección de datos de operaciones comerciales de granos realizadas por personas físicas y jurídicas (sin identificación directa), la creación de un nuevo archivo y el procesamiento de dichos datos (y otros) que permite encontrar patrones de comportamiento. Sin embargo, ningún dato implicará bajo ningún concepto la recopilación, archivo y tratamiento de datos sensibles y/o de carácter personal. Aun así, en todas R analizadas para el caso SIO-Granos se detectaron potenciales externalidades para con los dueños de los datos, los responsables de las bases y terceros, con algunos casos donde puede ser algo ambiguo. Así, la problemática excede a detectar datos personales y/o sensibles y presenta una complejidad y alcance que debe ser tenida en cuenta al momento de plantear cualquier iniciativa “*Big data*”. En ese sentido, si bien el contexto legal local cuenta con marco legal vigente específico, no siempre es posible anticiparse ante las diferentes externalidades exploradas, en este caso con la metodología de las 4R para el caso SIO-Granos.

Es menester dedicar atención a la problematización del “*Big data*”, dado que brinda la oportunidad de plantear dichas iniciativas considerando no solo las potencialidades sino también las posibles externalidades para con los dueños de los datos, los terceros y el resto de la sociedad y organizaciones (públicas y privadas). De este modo, es posible ir construyendo soluciones y servicios con nuevas tecnologías y procesos con real impacto en los resultados y cuidando en el proceso la integridad de los ciudadanos y organizaciones.

REFERENCIAS BIBLIOGRAFICAS

- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, communication & society*, 15(5), 662-679.
- Bronson, K., & Knezevic, I. (2016). Big Data in food and agriculture. *Big Data & Society*, 3(1), 2053951716648174.
- Brussa, V., Castillo, V., Ron, J., & Gerke, J. (2017). Social Big Data: desafíos en los gobiernos abiertos y los procesos colaborativos en la construcción de los datos. Paper presented at the Simposio Argentino sobre Tecnología y Sociedad (STS)-JAIIO 46 (Córdoba, 2017).
- Cuquet, M., Vega-Gorgojo, G., Lammerant, H., & Finn, R. (2017). Societal impacts of big data: challenges and opportunities in Europe. arXiv preprint arXiv:1704.03361.
- Disterer, G. (2013). ISO/IEC 27000, 27001 and 27002 for information security management. *Journal of Information Security*, 4(02), 92.
- Garcia-Fronti, J. (2013). *Nanotechnology Markets Design*.
- Kennedy, H., & Moss, G. (2015). Known or knowing publics? Social media data mining and the question of public agency. *Big Data & Society*, 2(2), 2053951715611145.
- Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in human geography*, 3(3), 262-267.

- Lam, F. (2010). Los sistemas de información de mercados agrícolas (SIMA) en las Américas. desarrollo de los agronegocios y la agroindustria rural en América Latina y el Caribe: conceptos, instrumentos y casos de cooperación técnica.(Eds.) IICA. San José, Costa Rica.
- Loreti, D. (1995). El derecho a la información: Paidós.
- McQueen, R., Garner, S., Nevill-Manning, C., & Witten, L. (1995). Applying machine learning to agricultural data. *Computers and electronics in agriculture*, 12(4), 275-293.
- Nissenbaum, H. (2011). A contextual approach to privacy online. *Daedalus*, 140(4), 32-48.
- Riva, F., Abbatemarco, M., & Cervino, A. (2017). El tratamiento masivo de datos en redes sociales virtuales: retos legales, éticos y de responsabilidad social. Paper presented at the Simposio Argentino de Informática y Derecho (SID)-JAIIO 46 (Córdoba, 2017).
- Scanavino, F. (2012). Derecho a la intimidad vs Derecho a la información. Antagonismo o complementariedad. *Sistema Argentino de Información Jurídica*.
- Schmarzo, B. (2013). *Big Data: Understanding how data powers big business*: John Wiley & Sons.
- Steinmann, M., Matei, S., & Collmann, J. (2016). A theoretical framework for ethical reflection in big data research *Ethical Reasoning in Big Data* (pp. 11-27): Springer.
- Transition, C. f. C.-o. w. E. E. i. (1993). Glossary of industrial organisation economics and competition law. Organization for Economic.
- Witten, L., Frank, E., Hall, M., & Pal, C. (2016). *Data Mining: Practical machine learning tools and techniques*: Morgan Kaufmann.