Evaluación de Impacto para la Protección de la Privacidad de Grandes Datos/Big Data

Mg. Abogada. María del Carmen Becerra[1] , Prog. Univ. Pedro Zarate[2]
Mg. Lic. María Claudia Gomez[3]

[1,2] Instituto y Departamento de Informática – FCEFN-UNSJ
[3] Departamento de Informática FCEFN-UNSJ
mcbecerra2008@gmail.com, pzarate@iinfo.unsj.edu.ar,
cacugomez@yahoo.com.ar

Abstract—En este trabajo se analiza el impacto en la gestión de privacidad de Grandes Datos/Big Data ante la entrada en vigor del nuevo Reglamento de la Unión Europea, en el marco de la Legislación Argentina y su futura reforma. Se basa en un modelo donde se identificaron el conjunto de prácticas y recursos de gestión, resultantes de la integración de guías, normas, estándares y obligaciones contractuales más relevantes para mantener a salvo los datos personales y generar confianza en el entorno. Se identificaron los riesgos y se exploran soluciones para desarrollar controles legales que se integren en un modelo que permita establecer los pasos para que cualquier tipo de organización pública o privada, pueda verificar el impacto organizacional de sus productos, procesos o servicio en el cumplimiento de sus objetivos estratégicos. Se establecen criterios para evaluar el impacto sobre la privacidad para las empresas según el tipo de dato que traten y el tipo de tratamiento que realicen conforme el imperativo legal.

Keywords: **Big Data/Open Data. Titularidad de los Datos, Integración de Estándares, Privacidad, Evaluación de impacto.**

1 Introducción

La Unión Europa ha publicado el 27/04/2016, el Reglamento sobre la Protección de Datos (RGPDUE), en él se incluyen conceptos como la privacidad desde el diseño que exige que la privacidad se tome en consideración desde la fase inicial, es decir desde el mismo diseño del producto o servicio, con ello no solo se conseguirá una mayor eficiencia en la protección de los derechos de los afectados, sino que la inclusión de estos conceptos obligará a las empresas a actualizar sus procesos internos para adaptarlos a estos requerimientos. Además a partir de la aplicación de este Reglamento desde mayo del 2018 será obligatorio el informe sobre el impacto de privacidad en la intimidad[1].Su alcance y exigibilidad se hará extensivo no solo a las empresas europeas sino también a las extranjeras que tengan acceso a datos, bien porque sean estos enviados desde Europa, bien porque dirija n sus servicios hacia personas físicas ubicadas en alguno de los países miembros de la Unión Europea.

[1]www.agpd.es/portalwebAGPD/temas/reglamento/index-ides-idphp.php

adfa, p. 1, 2011.

El fenómeno conocido como Big Data es el conjunto de tecnologías que permiten tratar cantidades masivas de datos provenientes de fuentes dispares, con el objetivo de poder otórgales una utilidad que les proporcione valor. Es un nuevo paradigma, que alude al enorme crecimiento en el acceso y uso de la información automatizada. La gestión y análisis de grandes datos generados por el uso de tecnologías digitales (como la telefonía móvil, las transacciones electrónicas y las redes sociales) son instrumentos eficaces para innovar en la gestión empresarial, la prestación de servicios públicos y el diseño e implementación de políticas de desarrollo[2].
En nuestro país se está analizando la reforma de la Ley 25.326, que coloca una responsabilidad proactiva en el prestador de servicios de tratamiento de datos, obligándoles a adoptar medidas para el cumplimiento de la Ley[3].
En este trabajo centraremos el objeto de investigación en la determinación de criterios para la Evaluación del impacto que produce Big Data sobre la protección de datos personales, e identificaremos los concretos tratamientos que se llevan a cabo sobre los mismos, basados en un modelo que permita la integración de estándares más relevantes para mantener a salvo la Privacidad. Se parte de la Evaluación del Impacto sobre la Protección de datos efectuado por el nuevo Reglamento de la Unión Europea y la guía que fija los criterios para su medición, para confrontarlo con el Proyecto de Ley Argentino, que modifica la Ley 25.326, a fin de establecer pautas comunes que permitan establecer nuevos criterios para evaluar la privacidad de los datos, logrando una gestión eficiente y eficaz de los datos generados y tratados por los servicios públicos, servicios sociales y servicios financieros.

2 Big Data/Open Data

Con el tamaño y complejidad de Big Data [1] llegan una infinidad de retos en materia legal y normativa, surgen así una serie de interrogantes ¿De quién son mis datos? ¿Por qué las empresas usan mis datos?, ¿De quién son esos enormes volúmenes de datos?, etc., sin embargo la respuesta no es unívoca, la titularidad de los datos presenta muchas aristas según como se los genere y se los trate bajo estrictas medidas de seguridad. Los usuarios regalan sus datos de consumo y localizaciones sobre todo por una vida más cómoda y simple. El negocio de los datos ya no se detiene, manipular a escala masiva y precisa las conductas y operar sobre nuestras emociones parece ser el gran reto. Según reporte de CISCO el trafico IP global anual alcanzara 3,3 ZB (ZB; 1000 Exabytes [EB] para el año 2021[4].
Así hoy más que nunca, la creciente y enorme cantidad de datos, del orden de ZB, generados por aplicaciones empresariales y de gobierno electrónico ponen en riesgo la privacidad en la nube y determinan la necesidad de un modelo de Integración para la gestión de seguridad de esos enormes conjuntos de datos [2].

---

[2]Arce Iván. Seguridad Tic. Desafíos y oportunidades para emprendimiento de base tecnológica en Argentina. www.fundacionsadosky.org.ar

[3]https://www.justicia2020.gob.ar/wp-content/uploads/2017/02/Anteproyecto-de-ley-PDP.pdf

[4]https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.html

Tal cual lo expresan Kenneth Neil Cukier and Viktor Mayer Schonberger, los grandes datos empiezan con el hecho que hay mucha información en estos días y hoy más que nunca está disponible para usos extraordinarios. Los grandes datos son más que solo comunicación: La idea es que podemos aprender de un gran cuerpo de información cosas que no podemos comprender cuando usamos solo pequeñas cantidades. Aunque existen varias definiciones de Big Data, a los efectos de este trabajo lo conceptualizaremos [1], como los enormes volúmenes de datos (estructurados, semiestructurados y no-estructurados) del orden de exabytes 10 a 18 bytes), cuyo almacenamiento y análisis (eg) análisis textual de mensajes de correo, Tweets, blogs) se puede hacer mediante base de datos especializadas que sirven para analizar esos enormes volúmenes de datos y conseguir así información y conocimiento.

Tal cual lo expresa Gabriel Baum"...es, en el sector de tecnologías de la información y la comunicación, una referencia a los sistemas que manipulan grandes conjuntos de datos (o data sets)". Las dificultades más habituales en estos casos se centran en la captura, el almacenado, búsqueda, compartición, análisis y visualización. La tendencia a manipular ingentes cantidades de datos se debe a la necesidad en muchos casos de incluir los datos relacionados del análisis en un gran conjunto de datos relacionados, tal es el ejemplo de los análisis de negocio, los datos de enfermedades infecciosas, o la lucha contra el crimen organizado.

Los científicos con cierta regularidad encuentran limitaciones debido a la gran cantidad de datos en ciertas áreas, tales como la meteorología, la genómica, la conectó mica, las complejas simulaciones de procesos físicos, y las investigaciones relacionadas con los procesos biológicos y ambientales, las limitaciones también afectan a los motores de búsqueda en internet, a los sistemas de finanzas y a la informática de negocios. Las líneas de investigación se orientan para el desarrollo y la Innovación en la Ciencia de los Datos, que se ocupa hoy más que nunca del análisis, interpretación y toma de decisiones estratégicas[5].

El fenómeno de Big Data constituye un fenómeno global que puede llegar a tener un impacto económico real y potencial, beneficiando tanto al sector público y privado, en el aumento de la productividad, la competitividad sectorial y la calidad de vida de los ciudadanos. En Argentina existe una alianza por los grandes datos entre el MINCYT y la Fundación Sadosky, que presentaron una estrategia de utilización de grandes datos para Argentina en el periodo 2013-2018. La Ley 26899 (B.O.09/12/2013) legisló en Argentina sobre los repositorios de datos abiertos que constituyen un modo de registro con normas internacionalmente compatibles de representación, búsqueda y acceso por datos personales, institucionales y temáticos, así como con condiciones de preservación a largo plazo y establece la responsabilidad de los usuarios, generadores y curadores de datos. Posteriormente en la plataforma de justicia 2020 se sentaron las bases para el proyecto de reforma de la Ley 25.326.

Big Data surge así como una nueva fase del paradigma intensivo en información y comunicación que abarca no sólo su dimensión tecnológica, sino también una dimensión social, económica, política y cultural [3]. Los big data (BD), como parte de la denominada internet de las cosas (IoT, por sus siglas en Ingles), deben ser considerados un fenómeno socio tecnológico en tanto que no solo están transformando la cultu-

[5]http://sedici.unlp.edu.ar/bitstream/handle/10915/52853/Documento_completo.pdf-PDFA.pdf?sequence=1

ra de la comunicación entre seres humanos, sino que intervienen simultáneamente en los espacios público y privado de los agentes haciendo que la frontera entre ambos se muestre más difusa que nunca antes [4].

Según la Comisión Económica para América Latina y el Caribe (CEPAL) "La gestión y análisis de esta enorme cantidad de datos digitales genera información valiosa para apoyar las preocupaciones emergentes que pueden ser de gran importancia para el desarrollo global", también lo establece en su agenda 2030 para el desarrollo sostenible[6].Estas nuevas oportunidades del Big Data van incrementando los riesgos y quizá el más relevante sea el que representa para la privacidad de las personas, con el tamaño y la complejidad también surgen una infinidad de retos en materia legal y normativa [2].

3 Titularidad de los datos

En su trabajo el Foro Económico mundial propone que a pesar de que los datos se refieran a individuos, se crean a partir de la interacción de diversas partes, de forma que todos ellos deben tener derechos y responsabilidades sobre esa información, los derechos deben ser comunes, no exclusivos [2]. Para responder al primer interrogante ¿de quién son los datos?, se debería esbozar que el titular de los recursos de computación debe poder probar de alguna manera que ha tomado las cautelas necesarias para garantizar la seguridad de sus sistemas en el ambiente de la integridad, confidencialidad y disponibilidad. Las normas de protección de datos aseguran que los individuos son quienes tienen el control sobre la información que de ellos se incorporan a las bases de datos, el foco de atención ya no debe ser el momento de la recolección de datos, sino el momento de la utilización de sus datos.

¿Respecto del Segundo interrogante planteado en el sentido de porqué las empresas utilizan mis datos?, la Legislación Argentina solo deja claro que es la empresa la que pone los medios para elaborar una base de datos, y quien tiene una serie de derechos sobre los mismos. Se debería hablar en términos de titularidad o más bien de uso en base a lo establecido en determinado acuerdo. Dependiendo del tipo de dato del que se trate el criterio de acceso que se aplicara es disímil. "La identidad de estos comsumers"[4], está íntimamente relacionada con la cultura de su organización y ligada inevitablemente al concepto de reputación. La dimensión, el tipo o el impacto de estos riesgos son difíciles de establecer a priori. ¿Cuáles son los riesgos potenciales si las consecuencias de una acción no son diagnosticadas ni atendidas? ¿Cómo podría una organización argumentar racionalmente sobre la justificación para correlacionar la información sobre el historial sanitario de una persona con información acerca de sus búsquedas en línea? ¿De qué modo afecta el resultado de esta correlación a nuestra identidad? [4].

El proyecto de reforma de la Ley de Protección de Datos Argentina establece que entre los principales factores de riesgo se pueden citar:

A-La evaluación sistemática y exhaustiva de aspectos personales de personas físicas que se base en un tratamiento de datos automatizado.

[6]https://biblioguias.cepal.org/c.php?g=447204&p=3192634

Las empresas líderes del sector han analizado en recientes publicaciones las amenazas y riegos del Cloud Computing y han efectuado recomendaciones. Sus preocupaciones se traslucen en aspectos como la gestión de datos, fundamentalmente en la propiedad de los mismos y la forma de operarlos y de tratarlos por parte de los proveedores, y en la identificación y control de acceso a los recursos. Además, se incluyen refuerzos de seguridad en cuanto a la confidencialidad de los datos [7].

El Open Cloud Manifiesto establece una serie de principios para garantizar una nube abierta. Colaboración abierta y un adecuado uso de los estándares para hacer frente a los retos que ofrecen la implantación de la nube. Los proveedores de servicios no deben retener a los usuarios en determinadas plataformas impidiendo la libertad de elección. Cuando sea pertinente los proveedores deben utilizar los estándares vigentes. Los nuevos estándares que se adopten deben promover la innovación y en ningún caso restringirla [8].

Los líderes de datos y análisis ahora vienen de todas las partes del negocio. Los programas de datos y análisis ya no son dirigidos exclusivamente por TI, sino que son creados por y para otros líderes empresariales. La naturaleza integrada, conectada y en tiempo real del negocio digital requiere la colaboración entre unidades organizativas históricamente independientes, y los programas de datos y análisis están en el centro de todo. Para darse cuenta de su visión de negocio digital y para la colaboración entre organizaciones, los negocios y las TI deben trabajar juntos para establecer una nueva visión estratégica, con roles y métricas [9].

Tal cual lo afirmado “las organizaciones que conformen alianzas para implementar una infraestructura Cloud comunitario deberán tener objetivos similares y un marco de seguridad y privacidad común para resguardar la información corporativa, en este sentido la Cloud Security Alliance asiste a las organizaciones en la toma de decisiones y en la adopción de estrategias”.

El proyecto de reforma de la Ley 25.326 establece que el tratamiento de datos que utilizan servicios de computación en la nube solo está permitido cuando garanticen los principios y obligaciones establecidos en la Ley. Además, establece una serie de responsabilidades que pone a cargo del Responsable del Tratamiento de los datos.

B- El Consejo de Europa ha señalado que los datos médicos forman parte de la esfera de intimidad de las personas, de manera que su trasmisión o divulgación solamente se pueden hacer en temas y problemas muy concretos y restringidos. Al respecto la Convención para la protección de los individuos en relación al tratamiento autorizado de datos personales, en su Art. 6 dispone que los datos personales relativos a la salud no pueden ser procesados automatizadamente, a menos que el ordenamiento nacional proporcione medidas de seguridad apropiados. Criterio que también recepta la Directiva 680/2016 del Parlamento Europeo y del Consejo relativo a los datos de las personas físicas destinados a los ámbitos policiales y a la justicia. El Reglamento de la Unión Europea sobre la protección de datos publicado el 27/04/2016, estableció que el consentimiento debe ser “Libre, específico, informado e inequívoco”. En el mismo sentido se pronuncia el proyecto de reforma de la Ley 25.326.

La Ley 25.326 ha brindado un estándar de puerto seguro al establecer que los datos vinculados a la salud solo podrán ser tratados, a fin de realizar oferta de bienes y servicios, cuando hubiesen sido obtenidos de acuerdo con los principios rectores establecidos por la misma ley y siempre que no causen discriminación. Estos datos no podrán ser transferidos a terceros sin el consentimiento previo, expreso e informado del

titular de los datos quien deberá recibir una noticia clara del carácter sensible de los datos y de que no está obligado a suministrarlos, junto con su derecho de solicitar su retiro de la base de datos. Además, estableció que "Toda actividad medico asistencial tendiente a obtener, clasificar, utilizar, administrar, custodiar y trasmitir información y documentación del paciente debe observar el estricto respeto por la dignidad humana y la autonomía de voluntad, así como el debido resguardo de la intimidad del mismo y la confidencialidad de sus datos sensibles". Esto último se ve acentuado después de la reforma del Código Civil y Comercial y especialmente por los principios establecidos en el Art. 9(Buena Fe), Abuso de derecho (10) Abuso de la posición dominante (Ar11), Orden Público (Art12) y Prohibición de renuncia de las Leyes.

El proyecto de reforma establece que una excepción cuando el tratamiento sea efectuado por los establecimientos sanitarios públicos o privados y los profesionales vinculados a la ciencia de la salud que traten los datos recogidos de fuentes de acceso restringido o de los datos de los pacientes que estén o hubieran estado bajo tratamiento de aquellos, respetando los principios del secreto Profesional. Además, hay que tener en cuenta que la Ley sobre derechos del paciente Nº26529 define en su art.5 el consentimiento informado" como la declaración de voluntad suficiente efectuada por el paciente, o por sus representantes legales en su caso, emitida luego de recibir por parte del profesional interviniente información clara, precisa y adecuada" y Nº26742, Decreto 1089/2012 y Art. 59 del CCC. En concordancia con ello hay que tener en cuenta las Normas sobre Buenas prácticas Clínicas (GCP-ICH), la declaración de Helsinki y las disposiciones del ANMAT Nº5330/97, 690/05,1067/08,6550/08 Res BNº1490/07, La Ley 11044 y Decreto 3385/08 y lo Establecido por el Ministerio de Salud mediante Decreto 3385/08 sobre consentimiento informado. Finalmente mencionaremos la Directiva 680/2016, del Parlamento Europeo y del Consejo relativo a las personas físicas destinados a los ámbitos policiales y a la justicia.

El tratamiento de datos sensibles a gran escala, o de datos relativos a antecedentes penales o contravencionales.

C- Tratamiento de datos mediante tecnologías que se consideren potencialmente invasivas de la privacidad.

La proliferación de dispositivos de video vigilancia, o los que utilizan a modo de soporte el propio cuerpo humano, como la biometría, y cuyo uso se extiende a terrenos cada vez más cotidianos y diversos como los utilizados para defensa, seguridad e inmigración utilizados en las áreas de gobierno [5]. La captación de imágenes en la vía pública, no se agota en la conducta de algunos ciudadanos que captan imágenes con sus cámaras, sino que también abarca la utilización de imágenes captadas por empresas de seguridad privada y las pertenecientes a la vigilancia estatal, a las que también se suma la enorme cantidad de imágenes que genera la proliferación de drones. El Proyecto de reforma de la Ley 25316 cita a estos como: La vigilancia electrónica, la minería de datos, el tratamiento a Gran Escala y el Internet de las Cosas.

La geolocalización para efectuar el seguimiento de los movimientos de los compradores y dado que las señales de GPS no suelen percibir en el interior de las superficies de venta o ciertos comerciantes minoristas pueda utilizar otras tecnologías electrónicas. Estas tecnologías incluir RFID en los productos o en los carros de compra, cámaras de video y varias tecnologías innovadoras aprovechando los teléfonos móviles entre ellas el reconocimiento facial en los servicios en línea y móviles en Argentina y en resto de los países latinoamericanos no existe gran conciencia sobre el uso de datos

biométricos. Estas videocámaras pueden ser excelentes para la gestión del flujo de tráfico, pero difíciles de utilizar para el seguimiento de la conducta de los individuos que entra en conflicto con la ley de protección de datos, a no ser que las imágenes se distorsionen para impedir las identificaciones individuales [6].

D-Tratamiento significativo no incidental de datos de niñas, niños y adolescentes, o dirigido especialmente a tratar datos de los mismos. El proyecto de reforma de la ley 25.326 establece en su Art 40 (Evaluación de impacto relativa a la protección de datos personales).

La Convención de los Derechos del Niño establece en sus Artículos 3 y 16 los derechos a una vida privada e intimidad de las niñas, niños y adolescentes, el Art. 75 Inc. 22 de la C.N. y art. 3,10 y 22 de la Ley 26.061.

Desde la ratificación de la Convención sobre los Derechos del Niño en 1990, con Jerarquía Constitucional desde 1994, Argentina ha logrado importantes avances en la materia. Así, desde 2005, el país cuenta con una Ley Nacional de Protección Integral de los Derechos de las Niñas, Niños y Adolescentes (Ley 26.061). A su vez, nuestro país confirmó su compromiso con otros instrumentos normativos entre los que se destacan: Ley de Protección Contra La Violencia Familiar (24.417), la Ley de Creación del Programa Nacional de Salud Sexual y Procreación Responsable (Ley 25.673), la Ley de Migraciones (Ley 25.871), Ley del Programa Nacional de Educación Sexual Integral (Ley 26.150), la Ley de Educación Nacional (Ley 26.206), la Ley de Prevención y Sanción de la Trata de Personas y Asistencia a sus Víctimas (Ley 26.364) y la Ley de Prohibición del Trabajo Infantil y Protección del Trabajo Adolescente (Ley 26.390).

4 Integración de guías, normas y estándares

Se parte de realizar una búsqueda en los documentos publicados hasta el momento, se ve que el proceso de autenticación de identificación biométrica se basa en las normas ISO/IEC 17.799, 27.001 y 27.002 (estándares centrados en la seguridad de la Información), COBIT (estándares centrados en la gestión), ITIL (estándares centrados en los organismos públicos), y ANSI NIST-ITL 378 (estándares centrados en la seguridad de datos biométricos).En las industrias de la salud, servicios públicos y servicios financieros existen muchas guías de diversos reguladores, quizá no obligatorias, pero si altamente recomendadas. También existen estándares contractuales como el PCI y DDS que controlan la información de transacciones con tarjetas de crédito. Por último, existen estudios de la industria para el seguimiento de la información publicada por organismos como el CERT (Computer Emergency Response Team) y las familias de estándares ISO [10].

Las Normas IRAM ISO/IEC 27.001 y 27.002[11], definen y documentan, los roles y responsabilidades de la seguridad de los empleados, contratistas y terceros en concordancia con la política de seguridad de la información de la organización. Garantizan la protección y privacidad de los datos según lo requieran las legislaciones y si fueran aplicables, las clausulas relevantes contractuales. Existen además, normas como la ISO 24.760 e ISO 29.100, que proporcionan un marco de referencia de alto nivel para la protección de los datos personales, y regulan la Gestión de identidad y

Privacidad. Aportan a la gestión de privacidad las ISO29134 (Privacy Impact Assessment), e ISO 29151, ISO 29190/91 que presenta un modelo de evaluación de la capacidad en privacidad. La norma ISO / IEC 27018: 2014 establece objetivos de control comúnmente aceptados, controles y guías para implementar medidas para proteger la Información de Identificación Personal (PII) de acuerdo con los principios de privacidad en ISO / IEC 29100 para el entorno de computación en nube pública. En particular, especifica directrices basadas en ISO / IEC 27002, teniendo en cuenta los requisitos reglamentarios para la protección de las IP que podrían ser aplicables en el contexto del entorno de seguridad de la información de un proveedor de servicios públicos servicios en la nube. Es aplicable a todos los tipos y tamaños de organizaciones, incluyendo empresas públicas y privadas, entidades gubernamentales y organizaciones sin fines de lucro que proveen servicios de procesamiento de información como procesadores PII a través de Cloud Computing bajo contrato con otras organizaciones. Estas directrices también podrían ser pertinentes para las organizaciones que actúan como controladores de PII; Sin embargo, los controladores de PII pueden estar sujetos a leyes, reglamentos y obligaciones adicionales de protección de PII, que no se aplican a los procesadores PII. ISO / IEC 27018: 2014 no pretende cubrir tales obligaciones adicionales. COBIT Acrónimo de "Control Objectives for Information and Related Technology (Objetivos de Control para la Información y Tecnologías Relacionadas), es un estándar desarrollado por la Information Systems Audit and Control Foundation (ISACA). Se destaca el rol de COBIT 5 en la estrategia de seguridad y objetivos de control, en el nuevo marco para la gobernanza en TIC´S, y guías de COBIT 5 sobre seguridad y riesgo. Los Estándares ANSI/NIST-ITL son estándares que se aplican internacionalmente para proteger los datos biométricos. El propósito de esta norma es que un sistema biométrico dactilar pueda realizar procesos de verificación de identidad e identificación, empleando información biométrica proveniente de otros sistemas.

En el modelo de la selección de estándares y normas, se utiliza el método de estudio de comparación de los mismos, se crean plantillas para la comparación de las similitudes respecto a la gestión de la identidad y la privacidad [12]. El modelo aplico la evaluación de normas y estándares, que proporcionan una base sólida para el cumplimiento de los objetivos de la Organización, en cuanto a la seguridad de los datos personales. Se basó en los modelos que tienen más fortalezas en la relación a la gestión de capacidad de servicio de TI, son el modelo ITIL [9] y las Normas ISO/IEC 27.001 y 27.002 que permiten que la información y la tecnología relacionada se rijan y se gestionen de manera integral en toda la empresa [10].

Es un modelo de integración donde las normas y estándares a ser evaluados responden a un modelo general de valuación de los sistemas de información, como conjunto de elementos interrelacionados para lograr un objetivo específico. En el modelo general se evalúan las normas y estándares de seguridad, en primer lugar se detectó el estado del arte de las normas y estándares, luego se las comparó en función del uso e impacto de cara al ciudadano, de cara al empleado y de cara al usuario [13]. En el Modelo de integración se estándares se midió el impacto desde tres perspectivas y la eficacia mediante el cumplimiento del principal objetivo que es la seguridad. Este análisis se hace previo a su implementación, formulando criterios de evaluación que permiten ponderar su importancia en la protección de la privacidad [14].

5 Privacidad

El modelo de evaluación del impacto sobre la protección de datos es necesario para brindar a los responsables del tratamiento de datos pautas suficientemente específicas, útiles y claras, máxime cuando en nuestro país se está analizando la reforma de la Ley 25.326 que coloca una responsabilidad proactiva en el prestador de servicios de tratamiento de datos, con obligaciones de los responsables y encargados del tratamiento de datos para adoptar medidas para el cumplimiento de la Ley [13].

Se establece la protección de datos desde el diseño y por defecto (Fig 1). El responsable del tratamiento aplicara medidas tecnológicas y organizativas apropiadas tanto con anterioridad como durante el tratamiento de datos a fin de cumplir los principios y los derechos de los titulares de datos [14].

Las medidas serán adoptadas teniendo en cuenta el estado de la tecnología, los costos de la implementación y la naturaleza, ámbito, contexto y fines del tratamiento de datos, así como los riesgos que entraña el tratamiento para el derecho a la protección de los datos de sus titulares.

El responsable del tratamiento aplicara las medidas tecnológicas y organizativas apropiadas con miras a garantizar que, por defecto, solo sean objeto de tratamiento de datos aquellos datos personales que sean necesarios para cada uno de los fines del tratamiento. Esta obligación se aplicará a la cantidad y calidad de datos personales recogidos, a la extensión de su tratamiento, a su plazo de conservación y a su accesibilidad. Tales medidas garantizaran en particular, que, por defecto, los datos personales no sean accesibles sin la intervención de la persona, a un número indeterminado de personas humanas.

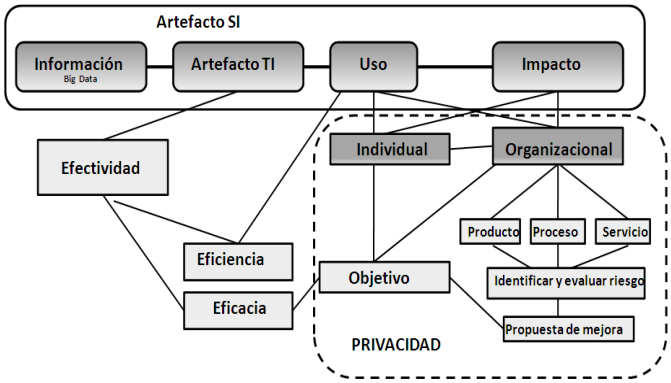

Fig. 1. Modelo de Privacidad

Se tienen en cuenta los criterios propuestos en otros informes y trabajos de la Agencia de Protección de datos española, GT29: 1) La relación entre el fin que se tomaren los datos y el fin del Modelo actual, 2) El contexto en el que se obtuvieron los datos y las expectativas de los sujetos, 3) La Naturaleza de los datos y su posible impacto, 4) Las

medidas de salvaguarda que se aplican como la imposición de condiciones de uso de los datos cuando estos se cedan a terceros. Además, se realiza un reconocimiento explícito del concepto de Privacidad por diseño y Privacidad por defecto[7]. También la solución que propone es la aplicación de un modelo de "ética de la privacidad". Son necesarios los mecanismos de legalidad internacional, por lo riesgos que el Big Data supone para la privacidad, esto ha hecho cambiar el foco ya no se debe cumplir solo con los principios de protección de datos ahora se debe tener en cuenta la privacidad desde el mismo modelo de negocio, es decir que la privacidad deja de ser un concepto legal para ser una prioridad del negocio [15].

Cada organización debe establecer su propia estructura de gestión y recoger en todos ellos las recomendaciones que resulten más útiles. El uso de estándares ayuda al cumplimiento de las leyes, reglamentos, acuerdos contractuales y políticos y a ganaren ventajas competitivas sobre otras organizaciones [16]. El aumento de la regulación y la legislación sobre la privacidad también está impactando en los entornos TI. Conforme las conclusiones del Proyecto referenciado la adopción de modelos y normas facilitan la rápida ejecución de los buenos procedimientos y ayuda a evitar retrasos innecesarios en el desarrollo de nuevos enfoques. Todas las empresas tienen que adaptar el uso de modelos y establecer normas para ajustar sus requisitos individuales [17].

El modelo genérico para la gestión de privacidad de grandes datos Big Data permitirá la definición de modelos más específicos según el contexto en que sean usados los datos [18]. La ventaja fundamental será que permitirá medir la efectividad y la eficiencia, teniendo en cuenta la eficacia y como se usa e impacta sobre la privacidad. [19].

6 Evaluación de Impacto

El proyecto de reforma de la Ley 25.326 establece la necesidad de una evaluación de impacto, para ello se deberán establecer criterios de impacto, pueden ser diversos según adoptemos alguna de las heurísticas definidas para la gestión de sistemas y sus criterios evaluables. La evaluación deberá incluir como mínimo: a) Una descripción sistemática de las operaciones de tratamiento de datos previstas y de los fines del tratamiento, inclusive, cuando proceda, el interés legítimo perseguido por el responsable del tratamiento; b) Una evaluación de la necesidad y la proporcionalidad de las operaciones de tratamiento de datos con respecto a su finalidad; c) Una evaluación de los riesgos para la protección de los datos personales de los titulares de los datos a que se refiere el apartado a); d) Las medidas previstas para afrontar los riesgos, incluidas garantías, medidas de seguridad y mecanismos que garanticen la protección de datos personales, y a demostrar la conformidad con la presente Ley, teniendo en cuenta los derechos e intereses legítimos de los interesados y de otras personas afectadas.

El Proyecto de Ley no incluye criterios para evaluar, por lo que en este trabajo se acude al estado del arte de las heurísticas definidas y sus criterios de evaluación. Si bien existen varias guías publicadas sobre el texto de la RGPDUE de cómo realizar

[7]https://www.masprivacidad.com/2017/03/01/la-privacidad-en-el- die%C3%B1o-y-por-defecto/

una evaluación de Impacto y obviamente el criterio de impacto es básico[8], se plantea el problema de la no graduación de sanciones. Los criterios para le evaluación de impacto son suministrados por SandaS GRC Privacidad versión RGPDUE. En las evaluaciones referidas se parte de identificar el escenario de riesgo y luego se calcula el riesgo inicial y se analizaran dos factores: Probabilidad e Impacto, se identificaran además los tratamientos que se llevaran a cabo sobre la información y además establece que se deberá preparar y planificar[9]. Para medir la probabilidad se suelen utilizar diferentes sistemas (porcentual, frecuencia diaria…)[10].

Este planteamiento puede enunciarse en base a los siguientes objetivos generales y específicos.

OBJETIVOS GENERALES. Para el desarrollo de esta investigación, se plantean los siguientes objetivos de índole general: Desarrollar un sistema de evaluación para medir el impacto en la privacidad de grandes datos mediante la revisión de heurísticas, aplicable al modelo genérico para la gestión de privacidad, estableciendo criterios que permitirán medir la efectividad y la eficiencia teniendo en cuenta la eficacia y como se usa en impacta sobre la privacidad

OBJETIVOS ESPECIFICOS. Los objetivos enumerados anteriormente se desarrollan en los que se relacionan a continuación:

Realizar un análisis y catalogación de métodos de evaluación basados en heurísticas.

Obtener una relación de heurísticas y subheurísticas relevantes para el proceso de evaluación del impacto sobre la privacidad.

Establecer una métrica para obtener el nivel de usabilidad.

Validar el método de evaluación propuesto mediante la realización de experimentos que permitan estimar la bondad del sistema de evaluación.

Desarrollar una herramienta que dé soporte al proceso de evaluación de impacto y permita almacenar los resultados en un formato estándar.

METODOLOGÍA DE INVESTIGACIÓN. La metodología de investigación seguida, desde el inicio de la investigación hasta la escritura de este trabajo se puede dividir en cuatro fases:

Fase 1: Estudio del estado del arte.

Fase 2: Planteamiento de los objetivos, contrastado y verificado por las conclusiones obtenidas tras el análisis del estado del arte.

Fase 3: Desarrollo de la investigación.

Fase 4: Diseño e implementación del prototipo de heurística.

Se ha tratado, además, de determinar, dentro de éstas, qué elementos de valoración proponen y cuales proporcionan una medida para establecer criterios, mediante una métrica asociada al proceso de evaluación, del nivel de privacidad de la información tratada. Respecto al área de la clasificación de seguridad de la información se han revisado varias propuestas atendiendo a diferentes parámetros de clasificación tratando de determinar si alguna de ellas se ajustaba a la requerida por la legislación vigente. Durante la segunda fase se han planteado los objetivos básicos de la investigación.

---

[8] www.govertis.com/como-realizar-una-eipd-en-proteccion-de-datos-rgpdue-parte-1

[9] www.govertis.com/parte-2-continuacion-como-realizar-una-eipd-rgpdue

[10] www.govertis.com/parte-3-como-realizar-una-eipd-rgpdue

Tras el estudio del estado del arte descrito anteriormente se decidió proponer un método de evaluación basado en heurísticas que no sólo sirviese de marco común de evaluación. La utilización de una herramienta Excel para calcular el impacto sobre la privacidad como soporte al sistema de evaluación propuesto agilizará en gran medida el desarrollo de ésta, la evaluación también podrá llevarse a cabo manualmente y por lo tanto, acometerse el proceso de validación empírica de la métrica propuesta. Para ello, antes y una vez desarrollada la herramienta, se realizarán varias evaluaciones en base al sistema SandaS, en el marco de varios experimentos. Tras esta última fase y durante el tiempo que dure la investigación, se llevaran a cabo revisiones y modificaciones de las propuestas iniciales hasta llegar a la propuesta de evaluación definitiva.
El propósito de esta iniciativa es redefinir parámetros existentes en los estándares internacionales actuales asociados con la Privacidad. El método de evaluación permite tener en consideración los aspectos que deberá cumplir el Responsable del Tratamiento de Datos. Este Método se basa en los criterios establecidos en la Ley.
A la hora de valorar criterios se establece una única relación de elementos a evaluar denominados criterios (sub heurísticos), agrupados en aspectos (heurísticas) que serán utilizados por todos los expertos implicados en el proceso de evaluación, así con una única relación de elementos a evaluar, se consigue unificar los criterios para llevar a cabo la evaluación del impacto sobre la privacidad. Se propone la siguiente escala de valoración: 1-Una escala de 0-10 que indica el grado de conformidad del evaluador con el cumplimiento del criterio. 2-Un valor textual, que indica si el criterio se cumple no. Este valor lo asigna el evaluador, pero, a efectos de cómputo, se aplica el mismo intervalo de medición (de 0-10). Se fijan así criterios sub heurísticos en función de la información tratada: Critica (SR) Problema es severo. Mayor (MA): Exige mucho esfuerzo para brindar una solución. Moderada (ME) Esfuerzo moderado para superar problema. El valor de relevancia de los aspectos matizara los resultados de la evaluación, de manera que, ante dos criterios de igual criticidad, será la relevancia de la información tratada la que determine cuál de ellos será prioritario en su arreglo (Tabla 1).
–NPD: No se cumple (valor 0)
-PDPSR Se cumple parcialmente con alto riesgo (valor de 2,5)
-PDPMA Se cumple parcialmente con riesgo medio (valor 5)
-PDPME Se cumple parcialmente con bajo riesgo (valor 7.5)
-SPD: Si se cumple (valor 10)
-NA: Criterio no aplicable
Si por ej. incluyéramos en una tabla heurística aspectos generales de Evaluación de Privacidad, descomponiendo la privacidad en sub-características y atributos medibles a los que se asocian métricas definidas genéricamente.

Aspectos Generales de la Evaluación		
CODIGO	CRITERIO	VALOR
PR.1	Cuando el tratamiento entrañe un alto riesgo a derechos fundamentales	0 1 2 3 4 5 6 7 8 9 10 NA
PR.2	Evaluación sistemática y exhaustiva de aspectos personales	0 1 2 3 4 5 6 7 8 9 10 NA
PR.3	Tratamiento de datos sensibles a gran escala	0 1 2 3 4 5 6 7 8 9 10 NA
PR.4	Tratamiento de datos mediante tecnologías invasivas de la privacidad	0 1 2 3 4 5 6 7 8 9 10 NA
PR.5	Tratamiento de datos de niños/as y adolescentes	0 1 2 3 4 5 6 7 8 9 10 NA
PR.6	Descripción del tratamiento de datos y sus fines	NPD PDPSR PDPMA PSPME SPD NA
PR.7	Evaluación de la necesidad de las operaciones del tratamiento de datos	NPD PDPSR PDPMA PSPME SPD NA
PR.8	Evaluación del riesgo para la protección de datos	NPD PDPSR PDPMA PSPME SPD NA
PR.9	Medidas previstas para afrontar los riesgos	NPD PDPSR PDPMA PSPME SPD NA

Tabla 1. Obligatoriedad sobre la Evaluación de Impacto

Los aspectos que se tienen en cuenta para un sistema común de evaluación con el objetivo de obtener una medida cuantitativa del Nivel de Privacidad y la indicación de la urgencia o prioridad en la mejora de los criterios en los que se ha detectado el riesgo se determinan en función del valor de severidad de los riesgos (Muy Alta-Alta-Media–Baja) y en función de los distintos tipos de tratamiento datos efectuados y el tipo de medida o salvaguarda adoptada (Tabla 2).

Ponderación de Aspectos	
Relevancia del Aspecto	Valor de Relevancia
Muy Alta	4
Alta	3
Media	2
Baja	1

Tabla 2. Ponderación de Aspectos

La presente investigación pretende contribuir a la mejora del contexto anterior proponiendo un método de Evaluación que hace uso de un modelo de usabilidad Web defi-

nido [20], teniendo en cuenta que la evaluación heurística es una revisión por parte de expertos en usabilidad a partir de unos principios establecidos por la disciplina IPO/HCI (Interacción Persona Ordenador) [20].
Sin embargo, como al principio de un proyecto, no siempre se van a conocer todas las posibles métricas que pueden llegar a interesar, es necesario, además definir qué métricas se desean calcular y determinar qué datos se van a recolectar. En etapas tempranas del proyecto se deben programar un conjunto de métricas básicas y recolectar una serie de datos. Cuando se realiza el plan de métricas se tienen en cuenta las instancias de evaluación de riesgos de modo que los datos estuvieran disponibles al momento de realizar estas actividades.
Actualmente se está trabajando en una herramienta Excel para calcular el impacto de privacidad conforme los criterios definidos y su desempeño, sus aspectos (y su descripción) y métricas para una evaluación de impacto de acuerdo a los estándares y según el modelo de privacidad propuesto en el Proyecto [15].
Además, se pretende la obtención de la medida cuantitativa del nivel de Privacidad y la indicación de la urgencia o la prioridad en la mejora de criterios en los que se ha detectado el fallo, se determinan valores de severidad de los aspectos y los criterios en función del tipo de información tratada, de esta forma para obtener un valor cuantitativo se debe considerar un concepto que incluya un factor de corrección y un cálculo del factor de corrección.

7 Conclusiones

En este trabajo se combina la experiencia de diferentes áreas del proyecto para dar información valiosa sobre lo que los grandes datos están haciendo, lo que puede hacer y lo que se debe permitir hacer. Esto significa un desafío potencial que enfrentarán las empresas para adaptarse al nuevo marco, donde los humanos deberán intervenir en las decisiones algorítmicas, para contribuir al proceso de toma de decisiones imponiendo requisitos al diseño de la privacidad.
El aumento de la regulación y la legislación sobre la privacidad también está impactando en los entornos TI. Con ellas, se pretende restringir la toma de decisiones individual automatizada, es decir los algoritmos que toman decisiones basadas en predicciones a nivel de usuario que afectan significativamente los mismos. La reforma de la legislación sobreprotección de datos planteara grandes desafíos para la industria, pero será una oportunidad para que los científicos de la computación tomen la iniciativa en el diseño de algoritmos y marcos de evaluación que eviten la discriminación y la explicación de sus Algoritmos mediante la adopción de criterios de trasparencia.
La Privacidad por diseño es una obligación legal que se impone a todo proyecto software, por lo que el aporte de este trabajo será establecer criterios para evaluar el impacto que provoca el tratamiento de datos personales en forma masiva y brindar pautas para la cuantificar el nivel de riesgo que implica el tratamiento para los datos personales de los individuos y de las empresas, siendo necesarios además establecer factores de corrección según las medidas tecnológicas y organizativas aplicadas por el responsable del tratamiento de datos, con miras a garantizar que la evaluación incluya como mínimo las pautas fijadas por Ley.

8 Referencias

[1] Joyanes, L. (2014). Big data: análisis de grandes volúmenes de datos en organizaciones. Barcelona: Marcombo. Ediciones Técnicas.
[2] Moreno, A., Redondo, T. (2016). Text Analytics: theconvergence of Big Data and Artificial Intelligence. End International. Journal of Interactive Multimedia and Artificial Intelligence.}
[3] Gil Gonzalez, Elena. Big Data-Privacidad y Protección de datos. Madrid 2016. ISBN978-84-340-2309-
[4] Colmarejo Fernández, Rosa. Una ética para Biga Data. Introducción a la gestión ética de datos masivos. Editorial UOC.01/01/2018.ISBN9788491169420.
[5] Bueres Alberto, Código Civil y Comercial de la Nación analizado, comparado y concordado. 1ra Ed. Bs.As. Ed. Ammurabi. ISBN 978-950-741-680-4
[6] Becerra. Navarro Mirta, Becerra, María del Carmen. Gestión Integral de Infraestructuras Criticas en las Organizaciones Locales alineados a las Normas IRAM ISSO 27.001 y 27002.
WSI - II Workshop de seguridad informática CACIQ 2013
[7] Neil, Robinson. Cloud: Understanding the Security, Privacy and Trust Challenges. RAND Corporation 2011.
[8] La gestión de la identidad digital: Una nueva habilidad informacional y digital. BID.
Universidad de Barcelona http://bid.ub.edu/24/giones2.htm
[9] http://blog.segu-info.com.ar/2012/07/como-se-construye-una-identidad-digital.
[10] Benantar, M.Access Control Systems-Security-Identity Management and trust Models New York:Springer.2006.
[11] AlleinniFéliz-Sánchez*, José Antonio Calvo-Manzano. Comparison of models and standards for implementing IT service capacity management.
www.redalyc.org/pdf/430/43038629008.pdf
[12] Becerra, María del Carmen, Zarate Pedro, Gómez Claudia. Modelo de integración de Estándares para la Gestión de Identidad y Privacidad. XXII Cacic2016. (WSI). Workshop de Seguridad Informática. UNS
[13] El derecho informático y la gestión de seguridad de la información una perspectiva con base a la norma ISO 27001. Revista del Derecho 2008. Biblioteca de Ciencia y Técnica de la
Nación
[14] Gonzalo Pérez-Tomé Estévez. Estudio sistemático de literatura de metodologías para la obtención de requisitos de privacidad.2015.www.dit.upm.es/…/TFM
[15]Becerra, María Del Carmen, Zarate Pedro, Gómez Claudia. Modelo Genérico para la Gestión de Privacidad de Grandes Datos/Big Data. 46 JAIIO. STS.2016
[16] DeLone, W, McLean, E., 2003. "Model of Information Systems Success: a ten years' update". Journal of Management
[17] Object Management Group, Inc. OMG. http://www.omg.org/2016.
[18] Jacobson, I; Booch, G; Rumbaugh, J. El proceso unificado del desarrollo de software. Pearson 2000.
[19] Diana M. Castillo Pinzón, DM. Enfoque para combinar e integrar la gestión de sistemas. 2010. ISBN 978-958-8585-06-2
[20] Sirius: Sistema de Evaluación de la Usabilidad Web Orientado al Usuario y basado en la Determinación de Tareas Críticas (PDF, 3.39 MB) de febrero de
2011.https://olgacarreras.blogspot.com.es/2011/.../sirius-nueva-sistema-para-la-evaluacion.html
[21] www:sedici.unlp.edu.ar/bitstream/handle/10915/23022/Documento_completo.pdf?
[22] https://eurolopd.com/Docs/general/Guia_EIPD.pdf