



UNIVERSITAT DE  
BARCELONA

**219**

**Big Data y el sector asegurador**

Estudio realizado por: Oscar Pallisa Gabriel  
Tutora: Montserrat Guillen

**Tesis del Máster en Dirección de Entidades  
Aseguradoras y Financieras**

Curso 2016/2017

Esta publicación ha sido posible gracias al patrocinio de ARAG SE, Sucursal en España



Cuadernos de Dirección Aseguradora es una colección de estudios que comprende las tesis realizadas por los alumnos del Máster en Dirección de Entidades Aseguradoras y Financieras de la Universidad de Barcelona desde su primera edición en el año 2003. La colección de estudios está dirigida y editada por el Dr. José Luis Pérez Torres, profesor honorífico de la Universidad de Barcelona, y la Dra. Mercedes Ayuso Gutiérrez, catedrática de la misma Universidad.

Esta tesis es propiedad del autor. No está permitida la reproducción total o parcial de este documento sin mencionar su fuente. El contenido de este documento es de exclusiva responsabilidad del autor, quien declara que no ha incurrido en plagio y que la totalidad de referencias a otros autores han sido expresadas en el texto.

## Presentación y agradecimientos

Para la realización de esta Tesis, el autor ha aprovechado su desempeño profesional para redactar un tema que se está desarrollando en la entidad aseguradora donde trabaja, Zurich España.

El BIG DATA ya se encuentra en un punto de madurez avanzado y esta práctica ya está ampliamente en uso en distintos sectores. Sin que el autor sea en absoluto un experto en el tema, sí que este tiene curiosidad y entusiasmo hacia esta materia, en la medida que gracias a la gestión masivas de datos, modelos de analítica avanzada, algorítmicos que mezclan conocimientos matemáticos, estadísticos o inteligencia artificial entre otros campos de conocimiento, sitúan a las tecnologías de la información en posición de avanzar al razonamiento humano.

Del mismo modo en el que ya tenemos costumbre a poder avanzar comportamientos meteorológicos, big data permite democratizar estas capacidades limitadas históricamente a unos pocos.

Agradecer a mi familia, a mi esposa Carol y mis hijos Adrià y Ariadna, que son quienes sufren mi plena dedicación profesional y consecuente merma en la dedicación familiar que se ha visto más acentuada aún en los últimos meses a propósito de este Máster. A mi padre Ramón, al que sin saberlo sigo sus pautas y a mi madre Gloria por su dedicación infinita y sin condición. Mis hermanas Lidia y Ester, por ser como son y concederme siempre el derecho a ser el que más se descuida en todo lo familiar. Siempre con ilusión, os quiero.

También agradecer a Oscar Gali, responsable de arquitectura de los sistemas de información y compañero de carrera, por sus enseñanzas en la materia y por hacer de mi desempeño profesional un gozo, al igual que a Cristina Rata, experta en analítica avanzada, que junto con su respectivos equipos han ayudado todos a llevar a cabo la puesta en marcha de las capacidades de analítica avanzada en Zurich España y cuyas experiencias se beneficia en buena parte el contenido de esta tesis.

Por último agradecer a Jordi Calbet, por haber pensado en mí para cursar este Máster que sin duda ha enriquecido mis conocimientos y ha saldado mis ansias de conocer más la industria aseguradora.



## Resumen

Esta tesis recoge una síntesis de cuáles son los orígenes y las dificultades de la gestión de grandes volúmenes de datos, cuáles han sido las limitaciones y avances tecnológicos para permitir la democratización de su gestión en lo que se ha bautizado generalmente con el término de *Big Data*. Se pretende ahondar sobre qué permite esta tecnología y cuáles son los ámbitos de utilización prácticos generales de analítica avanzada dentro del sector para posteriormente recorrer más detalladamente el proceso de desarrollo de un modelo predictivo, en particular un modelo de propensión de abandono de clientes.

**Palabras Clave:** Big Data, Modelos Predictivos, Analítica Avanzada, Modelo de propensión de fuga, grandes volúmenes de datos, ámbitos de aplicación.

## Resum

Aquesta tesis recull una petita síntesis de quins han estat els orígens i les dificultats de la gestió de grans volums de dades, quines han estat les limitacions i el avanços tecnològics que han permès la democratització de la seva gestió en el que s' ha batejat generalment amb el terme Big Data. Es pretén aprofundir sobre què permet aquesta tecnologia i quins son els àmbits d' utilització pràctics generalistes d' analítica avançada dintre del sector per posteriorment recórrer més detalladament el procés de desenvolupament d' un model predictiu, en particular un model de propensió de abandonament de clients.

**Paraules Clau:** Big Data, Models predictius, analítica avançada, model de abandonament de clients, Grans volums dades, àmbits aplicació

## Summary

This thesis presents a brief synthesis of the origins and difficulties of managing large volumes of data, which have been the limitations and technological advances to allow the democratization of its management in what has been generally defined as Big Data. It is intended to delve into what this technology enables and what the general practical areas of use of advanced analytics are within the sector, and then to further explore the process of developing a predictive model, in particular a model of customer churn.

**Keywords:** Big Data, Predictive Models, Advanced Analytics, Customer Churn, Large Data Volumes.



# Índice

|   |           |
|---|-----------|
| <b>1. PRESENTACIÓN DEL PROBLEMA</b> .....   | <b>9</b>  |
| <b>2. QUÉ CAUSA EL <i>BIG DATA</i></b> .....  | <b>11</b> |
| <b>3. LA RUPTURA DEL PARADIGMA</b> .....  | <b>15</b> |
| 3.1 LA GESTIÓN DE BASES DE DATOS RELACIONALES Y SUS LIMITACIONES .....                        | 16        |
| 3.2 REMEDIACIÓN A LAS LIMITACIONES DE PROCESAMIENTO VERTICAL .....                            | 19        |
| <b>4. ÁMBITOS DE APLICACIÓN</b> .....   | <b>21</b> |
| 4.1 ÁMBITOS DE APLICACIÓN EN EL SECTOR FINANCIERO .....                                       | 21        |
| 4.2 ANÁLISIS DEL ESTADO DE MADUREZ DE ANALÍTICA AVANZADA EN EL SECTOR ASEGURADOR              | 22        |
| 4.2.1 <i>Analítica en el ámbito de definición de producto</i> .....                           | 23        |
| 4.2.2 <i>Analítica en la suscripción</i> .....  | 24        |
| 4.2.3 <i>Analítica en la detección de fraude</i> .....  | 24        |
| 4.3 OTROS EJEMPLOS PRÁCTICOS .....  | 25        |
| 4.3.2 <i>Indexación documental</i> .....  | 25        |
| 4.3.3 <i>Scoring de impago</i> .....  | 26        |
| <b>5. MODELO DE PROPENSIÓN AL ABANDONO DE AUTO</b> .....                                      | <b>29</b> |
| 5.1 EL ENTENDIMIENTO DE NEGOCIO.....  | 29        |
| 5.1.1 <i>Definición del objetivo y criterio de éxito</i> .....                                | 29        |
| 5.1.2 <i>Evaluación de la situación y aproximación</i> .....                                  | 29        |
| 5.1.3 <i>Definición de los objetivos de minería de datos, y descripción del alcance</i>       | 30        |
| 5.2 ENTENDIMIENTO Y PREPARACIÓN DE LA INFORMACIÓN/DATOS .....                                 | 31        |
| 5.3 MODELIZACIÓN Y EVALUACIÓN.....  | 33        |
| 5.4 DESPLIEGUE.....   | 34        |
| <b>6. ORGANIZACIÓN</b> .....  | <b>35</b> |
| <b>7. AMENAZAS AL <i>BIG DATA</i>: REGLAMENTO GENERAL DE PROTECCIÓN DE DATOS (GDPR)</b> ..... | <b>37</b> |
| <b>8. CAPÍTULO DE CONCLUSIONES</b> .....  | <b>41</b> |
| <b>9. BIBLIOGRAFÍA</b> .....  | <b>43</b> |
| <b>OSCAR PALLISA GABRIEL</b> .....  | <b>45</b> |





# **Big Data en el sector asegurador**

## **Modelo de propensión al abandono de auto**

### **1. Presentación del problema**

El problema que plantea este ejercicio es el de desarrollar un modelo de predicción que de una manera lo más certera posible sea capaz de anticipar cuál es la probabilidad que un cliente de una entidad aseguradora renueve su póliza de automóvil con la entidad. El interés de este modelo radica en dotar a la empresa de esta información para poder discernir cuáles son las acciones más idóneas para cada escenario. Esto es, entre otras, el modular las actividades de incentivación como descuentos en renovación, limitación o dotación en campañas de marketing o simplemente abandono de cualquier actividad proactiva por parte de la entidad aseguradora.

La metodología que se seguirá es la de redactar un caso práctico real dentro de la empresa donde el autor ejerce su desarrollo profesional y para lo que se dotará de capacidades tecnológicas conocidas como Big Data. Se detallará cuál es el proceso de elaboración del mismo que recorre a grandes rasgos la identificación de las variables a utilizar, la extracción y manipulación de los datos de los sistemas de información de modo que se pueda modelizar para posteriormente integrar de manera recurrente en los procesos y sistemas de información.



## 2. Qué causa el **BIG DATA**

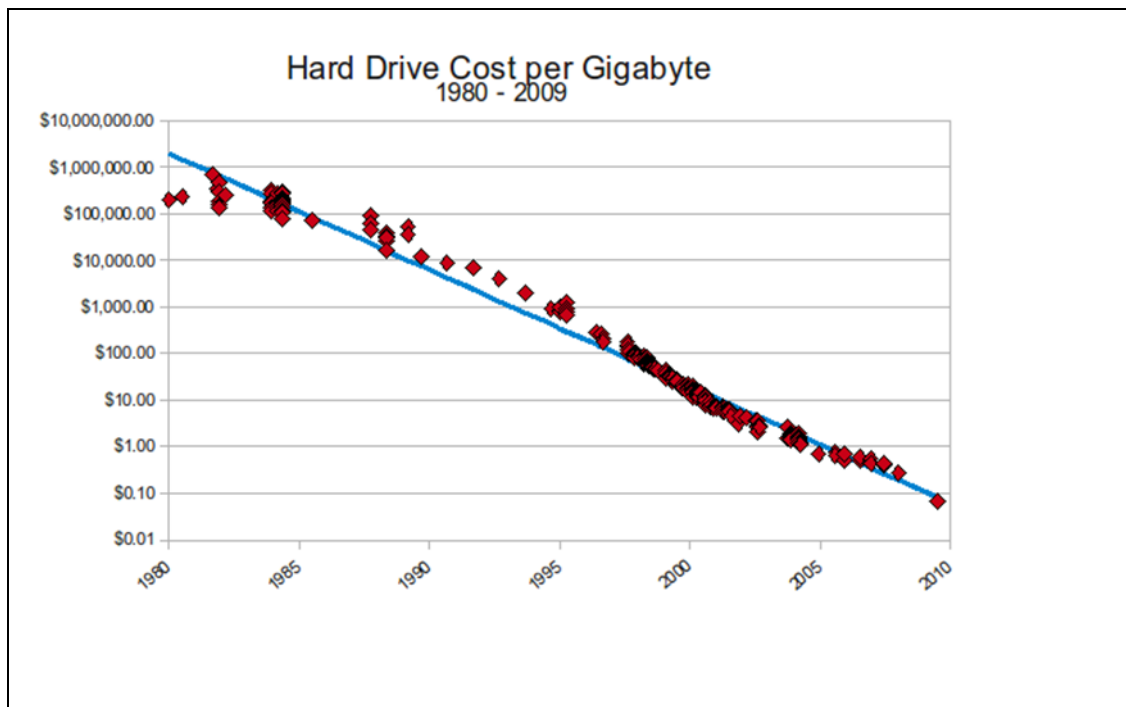
Antes de iniciar a describir cuáles han sido los avances técnicos que han producido la aparición de nuevos paradigmas en la ingestión y tratamiento de grandes volúmenes de datos conocidos como «Big Data», creo conveniente hacer un recorrido sobre las razones que han producido que ciertas entidades iniciaran investigaciones en esta materia y el porqué de la aparición de esta tecnología disruptiva.

Existe documentación de los años cuarenta donde se adelantaba la dificultad de la gestión de los volúmenes crecientes de información, con ejemplos que hacían referencia a la dificultad en gestionar el censo de Estados Unidos o los directorios de bibliotecas públicas entre otros. Si nos centramos en la historia más reciente desde la introducción de los ordenadores personales, cualquier persona de edad media en la actualidad puede recordar los incrementos de los volúmenes de almacenamiento de datos desde los inicios de la utilización de los sistemas de información en masa a finales de los años 80 hasta la fecha.

Los más antiguos recordarán tarjetas perforadas, otros quizás cintas de casete, algunos maduros iniciarían su andadura con disquetes de 5 ¼ o 3 ½. La generación *milenial* muy probablemente no reconocerá nada anterior a un pen-drive de múltiples GB de almacenamiento, discos sólidos o almacenamiento ilimitado en el *cloud*. La capacidad de almacenamiento entre los primeros dispositivos y los actuales de uso doméstico se ha multiplicado por 100 millones.

Pues bien, la diferencia en 30 años tanto de la capacidad de almacenaje como su coste asociado de fabricación se pueden ver reflejados en el gráfico a continuación. Para facilitar la interpretación al lector, podemos comparar el coste de 1GB de almacenamiento en 1980, que rondaba los 193.000 USD con apenas 0,07 USD en 2010. Esto muestra una correlación exponencial donde el almacenamiento por unidad de coste se ha doblado cada aproximadamente 14 meses.

Gráfico 1. Coste de almacenamiento



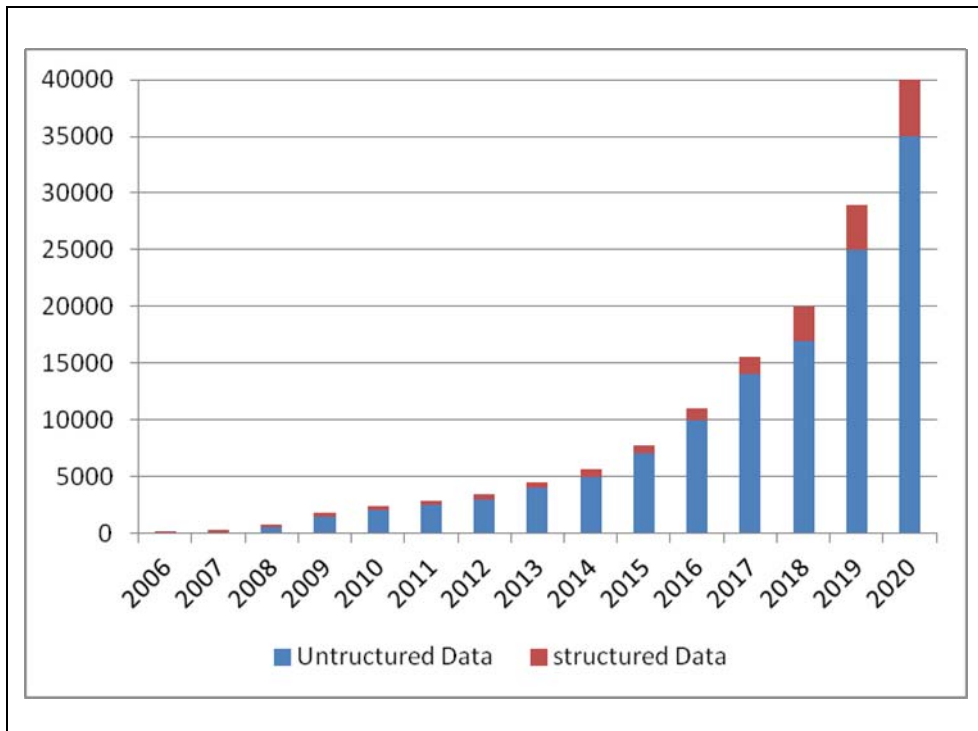
Fuente: Matthew Komorowski

Adicionalmente a los crecimientos de capacidad de almacenamiento y a su accesibilidad económica se le deben sumar otros factores que han actuado como factor multiplicador en la generación de datos, en la producción de lo que conocemos como «Big Data».

A los datos generados inicialmente por seres humanos se le añadieron las aplicaciones corporativas en masa, posteriormente se le sumó la web, donde individuos y corporaciones generaron mayores cantidades de información. En la actualidad, no solo estos sino multitud de sensores de todo tipo (el denominado internet de las cosas), desde cámaras, cajeros automáticos, sensores climatológicos, pulseras inteligentes y demás dispositivos capturan de manera imparable información que es posteriormente almacenada.

IBM argumenta que la sociedad genera 2,5 quintillones de bytes de datos anualmente de modo que el 90% de los datos existentes en la actualidad se han generado sólo en los dos últimos años.

Gráfico 2. Crecimiento de la generación de datos en exabyte (billones de GB)



Fuente: Patrik Cheesman



### 3. La ruptura del paradigma

En el capítulo anterior hemos visto que tanto la reducción exponencial de los costes de almacenamiento como el incremento de la generación de información han dado pie a la aparición de grandes volúmenes de datos conocidos como «Big Data».

Sin querer adentrarme en temas demasiado técnicos, me gustaría añadir un nuevo concepto para que uno entienda mejor las dificultades que conllevan la ingestión (léase almacenamiento) y tratamiento (léase procesamiento) de la información.

Tras la ingestión de los datos, normalmente existe un paso posterior que conlleva la gestión y tratamiento de los mismos. Habitualmente los datos se almacenan con un propósito que conlleva algún tipo de procesamiento de mayor o menor complejidad. Esto puede ser tan sencillo como un registro de datos de autenticación para acceder a una web hasta el tratamiento complejo de reconocimiento de imágenes en tiempo real, con multitud de algoritmos a procesar. En la imaginación está el límite.

En el ámbito de las tecnologías de la información, existe un término conocido como capacidad de procesamiento, que está supeditada muy fuertemente tanto por el procesador (CPU, Central Processor Unit) como la capacidad de transportar esos datos entre los discos de almacenamiento de información y los procesadores encargados de realizar los cálculos con ellos. Es decir, para procesar los datos, debo tener un procesador suficientemente potente y además, debo poder trasladar rápidamente los datos desde el dispositivo donde los almaceno hasta la central de procesamiento. Y esto se repite tantas veces como se requiera hasta conseguir el resultado deseado.

Es aquí donde existe un segundo elemento de vital importancia para el tema que nos ocupa. La necesidad de procesamiento y de transmisión de datos debe ser mayor cuanto mayor es el volumen de datos a tratar.

Imaginemos por un momento una cola de supermercado, donde cada cliente que espera en la línea del cajero/a simboliza un dato, y donde el cajero/a es el centro de procesamiento (CPU). Si el dueño del supermercado quiere que sus clientes esperen lo menos posible en su establecimiento tiene dos opciones, o cobrar más rápido mediante cajero/as más potentes (lo que se entiende como aumento de la capacidad vertical de procesamiento) o bien poner más cajero/as a atender a los clientes (lo que se conoce como procesamiento en paralelo u horizontal).

Estarán de acuerdo conmigo que estas opciones serán válidas siempre y cuando la cantidad de clientes (datos) que llegan a la caja sea inferior que la capacidad de procesamiento en vertical o en paralelo, de lo contrario no habrá otra opción que resignarse a la aparición de colas e interminables esperas. Lo que resultará en largos tiempos de procesamiento. Esto es lo que

ocurre cuando la cantidad de información es superior a la capacidad de procesamiento.

Con esto uno puede preguntarse, ¿desde cuándo el sector de las tecnologías de la información trata con grandes volúmenes de datos?. ¿Es esto un nuevo fenómeno acaecido en los últimos años?. La respuesta es que no, el término «Big Data» se empleó por primera vez en un artículo de los investigadores de la NASA que fecha Julio 1997, Michael Cox y David Ellsworth. Ambos afirmaron que el ritmo de crecimiento de los datos empezaba a ser un problema para los sistemas informáticos actuales.

El sector y más concretamente los centros de supercomputación hace décadas que tratan con este fenómeno de manera exitosa. Proyectos de investigación de predicción meteorológica, estudios de biomedicina sobre ADN entre otros requieren grandes volúmenes de datos y simulaciones. La diferencia está en que los centros de supercomputación son reducidos (debido a sus altos costes de adquisición de tecnología puntera y refresco constante) y accesibles solo a unos pocos, sin que estos puedan dar accesibilidad a las grandes masas en la gestión del «Big Data».

Entonces el lector puede preguntarse, ¿qué ha tenido que ocurrir para que exista una capacidad disponible de manera extensible en la sociedad, de tratar grandes cantidades de datos sin tener que recurrir a centros de supercomputación al alcance de unos pocos?. Para dar explicación a esto introduciré el tercer y último concepto que inició su andadura a partir de los motores de búsqueda de internet de Google, y para ello deberemos recurrir al concepto de las bases de datos relacionales y la evolución de este estándar.

### **3.1 La gestión de bases de datos relacionales y sus limitaciones**

Durante años el ámbito de las tecnologías de información ha ido evolucionando. Podemos decir que las tecnologías de la información y la computación en general han sufrido un avance exponencial si pensamos que es un campo que apenas cumple 50 años desde su tímida introducción mediante sistemas de computación centralizados en la empresa en los 60.

En el año 1970, Edgar F. Codd, matemático, trabajador de IBM publica un artículo en el que se explicaba la forma en la que podía accederse a la información almacenada en bases de datos de gran tamaño sin saber cómo estaba estructurada la información o dónde residía dentro de la base de datos. Hasta aquél momento, para recuperar la información se necesitaban conocimientos informáticos relativamente sofisticados, no existía un estándar entre los distintos sistemas y gestores de almacenamiento de datos, por lo que su manipulación era una tarea ardua que exigía tiempo y muchos recursos.

Esto dio pie a un estándar de estructuración, almacenamiento y acceso de datos conocidos como sistemas de gestión de bases de datos relacionales



(sus siglas en inglés *RDBMS relational data base management systems*) que han sido y siguen siendo el estándar mayoritario de almacenamiento de datos. Muy a groso modo, las bases de datos relacionales están compuestas por una serie de tablas inter-relacionadas, donde cada tabla tiene un número de columnas que definen la tipología de cada campo, y se extienden a modo de filas.

Existen varios proveedores de software de bases de datos relacionales que serán conocidas por el lector, tales como ORACLE (empresa tecnológica de mayor capitalización bursátil en el momento de la elaboración de este documento), IBM DB2 o Microsoft SQL Server entre otros. Todos estándares de mercado cuyo acceso a su información se realiza mediante un lenguaje estándar denominado SQL (de sus siglas en inglés Structured Query Language), lenguaje de consultas estructurado.

La gran mayoría de transacciones de datos actuales tales como utilizar tarjetas de crédito, acceder a cuentas bancarias, realizar reservas de viaje, realizar compras a través de Internet, utilizan estructuras basadas en la teoría de base de datos relacional mediante SQL.

Tras esta dosis de tecnicismos cabe explicar cuáles son las virtudes de las bases de datos relacionales y más detenidamente cuáles son sus limitaciones.

En el apartado de las virtudes comentaríamos la estandarización, facilidad de uso mediante una estructura basada en tablas, filas, columnas y las relaciones entre estas. Se le pueden añadir muchos otros, como la flexibilidad, seguridad mediante la autorización de accesos a tablas o la manipulación mediante un lenguaje casi algebraico como es el SQL.

En esta ocasión es relevante fijarse en las limitaciones, que se podrían resumir en tres grandes ámbitos que como descubrirá el lector, fueron el detonante del cambio de paradigma que daría pie a la eclosión del «Big Data».

Una de las desventajas fundamentales de las base de datos relacionales es la capacidad computacional que requieren, su rendimiento. Si el número de tablas y sus relaciones son extensas así como las mismas tablas (filas) se genera un efecto directo en el rendimiento de las consultas. Es decir, para poder procesar grandes cantidades de datos, se requiere una capacidad significativa de procesamiento y esta debe ser vertical.

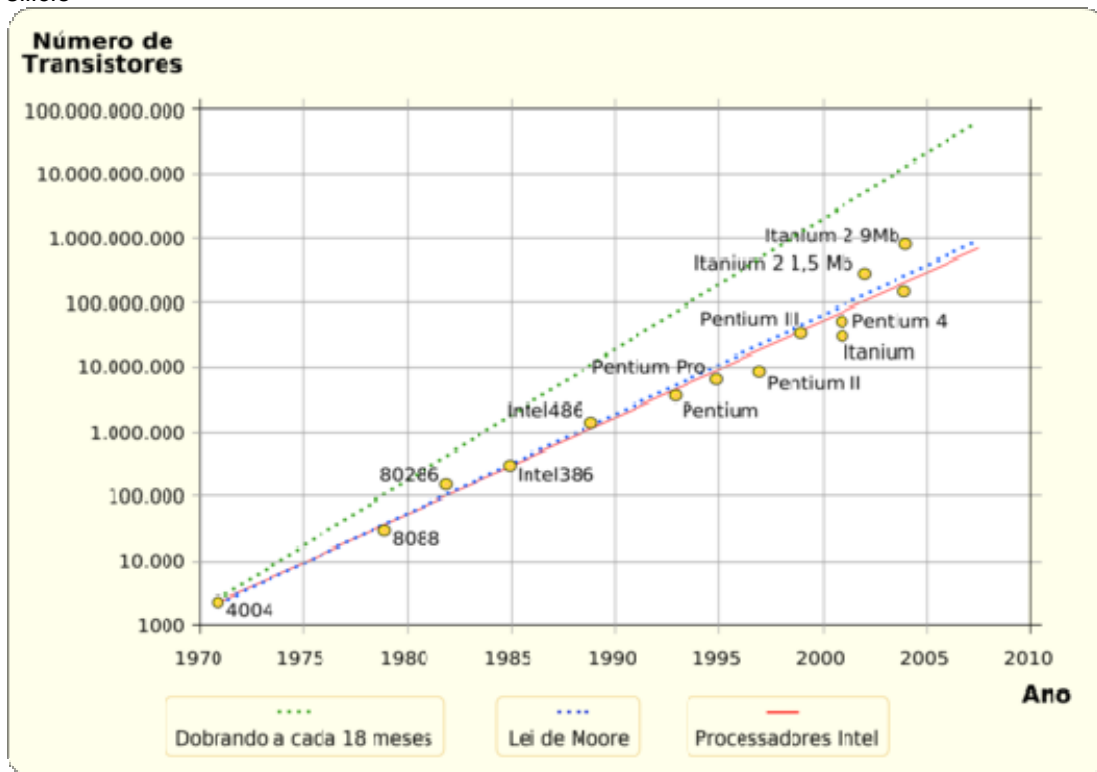
Utilizando el símil del supermercado de nuevo, donde los datos son los clientes y los procesadores los cajeros/as, podríamos decir que el estándar de base de datos relacionales tiene una limitación de rendimiento, para procesar más datos se requiere el aumento de la capacidad de la unidad de procesamiento, el cajero más rápido (lo que se denomina técnicamente, escalabilidad vertical), pero no permite de manera eficiente aumentar el número de cajeros en paralelo (lo que técnicamente se denomina escalabilidad horizontal).

Mientras que el avance en potencia computacional de los microprocesadores también ha experimentado una correlación similar a la de los datos, aumentando en potencia y decreciendo en costes, podemos decir que el aumento de grandes cadenas de datos superaba con creces las capacidades actuales de procesamiento para poder obtener resultados sin largos periodos de espera.

En los años 60, se formuló una ley empírica conocida como la Ley de Moore. Esta ley pronosticaba que el número de transistores en un procesador (que supedita su capacidad de cálculo) se doblaba cada dos años y que esta ley se cumpliría durante las siguientes dos décadas.

El problema y preocupación general ocurrió cuando el mismo Moore determinó en el año 2007, que esta ley “caducaría” pasados entre 10 o 15 años. Esto quiere decir que la tecnología actual de procesadores de silicio llegará a su máxima capacidad en breve. ¿Qué consecuencia tendría esto en nuestro ejercicio?, pues que no existiría expectativa en incrementar la capacidad de cálculo en nuestros procesadores para lidiar con los incrementos de datos a gestionar y nos condenaría a sufrir largas colas de tiempos de procesamiento, días, meses e incluso años.

Gráfico 3. Ley de Moore de crecimiento de capacidad computacional en procesadores de silicio



Fuente: José Maria Silveira Neto

Con la eclosión y crecimiento de los volúmenes de datos, la capacidad de procesamiento vertical no daban abasto mediante la capacidad computacio-

nal empresarial accesible, y sólo las grandes corporaciones o los centros de supercomputación podían acceder a estas.

Para los lectores optimistas, cabe decir que la tecnología avanza imperiosamente y se estima que nuevas tecnologías como el uso del grafeno u otros, permitan en un futuro cercano seguir creciendo en la capacidad vertical de procesamiento más allá de la tecnología actual de mercado aunque esto puede ser objeto de otras tesis en el ámbito de la ingeniería computacional.

En el próximo capítulo trataremos cómo se ha solventado esta diferencia entre la capacidad de procesamiento y el crecimiento exponencial de los datos mediante un cambio en el paradigma del almacenamiento y gestión de los datos. Avancemos un poco más.

### **3.2 Remediación a las limitaciones de procesamiento vertical**

En 2003, Google publica unos artículos explicando un nuevo método de almacenamiento de ficheros denominado Google File System y algo después MapReduce en 2004. Lo que daría lugar al estándar de Hadoop, a raíz de la necesidad de identificar sistemas nuevos para gestionar la explosión de datos de la web. De descarga gratuita, y libre para potenciarlo y mejorarlo. Hadoop se define como un método de código abierto para almacenar y procesar los datos que “permite el procesamiento en paralelo distribuido de enormes cantidades de datos en servidores estándar del sector, económicos, que almacenan y procesan los datos, y que pueden escalarse sin límite”

¿Qué virtudes tiene esta tecnología en palabras llanas? Nos permite superar dos escollos, por un lado permite tratar grandes volúmenes de datos mediante una estructuración de los datos distinta a la relacional tal y como se describe en el apartado 3.1. Y esto a su vez nos permite poder dotarnos de escalabilidad de procesamiento horizontal. Es decir, ya no necesito aumentar procesadores más potentes sino que puedo tener procesadores estándar pero en más cantidad trabajando en paralelo. Utilizando el caso del supermercado, solucionamos el problema de las colas mediante la extensión de más cajeros/as con productividad/potencia convencionales.

Por lo descrito anteriormente, podemos decir que la definición de nuevos estándares de almacenamiento de datos no relacionales ha dado pie a un nuevo paradigma en la gestión de los datos, de manera que lo que hasta la fecha requería capacidad computacional de procesamiento centralizada, vertical, supeditado a incrementos de potencia de procesadores con menor índice de crecimiento y mayor coste (y que llega a su final según la ley de Moore), ahora se puede suplantar por una serie de servidores corrientes, asequibles, integrados de manera paralela en línea para alcanzar los mismos resultados.

En otras palabras, para lo que anteriormente necesitábamos grandes ordenadores de coste elevado, ahora se puede lograr mediante bases de datos no-

relacionales, de código abierto (gratuitas) y servidores convencionales a precios asequibles.

Una vez el lector ha llegado hasta este punto creo estará en mejor condición para interpretar de forma más clara las definiciones del concepto BIG DATA que encontramos de distintas fuentes:

Big Data es un término aplicado a conjuntos de datos que superan la capacidad del software habitual para ser capturados, gestionados y procesados en un tiempo razonable. *WikiPedia*

Big Data es una colección de datos de fuentes tradicionales y digitales dentro y fuera de su organización, que representa una fuente para el descubrimiento y el análisis. *Forbes*

Big Data es un activo de información de gran volumen, alta velocidad y alta variedad que exige formas innovadoras de procesamiento de información rentables e innovadoras para una mejor comprensión y toma de decisiones. *Gartner*

## 4. Ámbitos de aplicación

En los capítulos anteriores se ha hecho un rápido recorrido tanto a los agentes que han contribuido al crecimiento de la aparición de grandes volúmenes de datos, cuáles han sido los estándares de gestión de datos de las bases de datos relacionales y cuáles han sido las tecnologías que han permitido romper las barreras que limitaban la gestión de estos grandes volúmenes de datos.

Una vez superado el escollo tecnológico, ¿qué nos queda?, pues si hay algo más importante que poder gestionar grandes volúmenes de datos es el saber qué hacer con ellos.

Este capítulo hace un pequeño recorrido de algunos casos de uso generalistas antes de adentrarse en aquellos que pudieran darse dentro del mundo empresarial para finalizar en el sector asegurador.

La utilización de grandes volúmenes de datos puede ser muy diversa. Después de “*hype*” que acompañó a este concepto desde sus inicios hasta la fecha, se han podido ver usos de muy diverso índole. Existen casos recientes locales ilustrativos tales como el del Ayuntamiento de Barcelona, que utilizando análisis de tráfico de población a partir de sus señales telefónicas de dispositivos móviles puede modelar la red de transporte público, a los más incipientes de finales de los 90 de la mano de Amazon, donde se utilizaban análisis masivos de información para recomendar lecturas, servicios o productos según un amplio espectro de variables.

En el caso del sector asegurador, y en comparación a otros sectores, considero que debería en principio partir de una posición más ventajosa, en el sentido que el sector lleva intrínseco en su esencia la ciencia actuarial, el uso de información masiva (léase teoría de los grandes números) para poder dimensionar los riesgos y sus costes de comercialización mediante sistemas de mayor o menor complejidad estadística.

### 4.1 Ámbitos de aplicación en el sector financiero

Mientras que los ámbitos de aplicación de *big data*, junto con las nuevas capacidades y competencias de los expertos en analítica avanzada puede ser muy amplio, en este capítulo describimos potenciales usos.

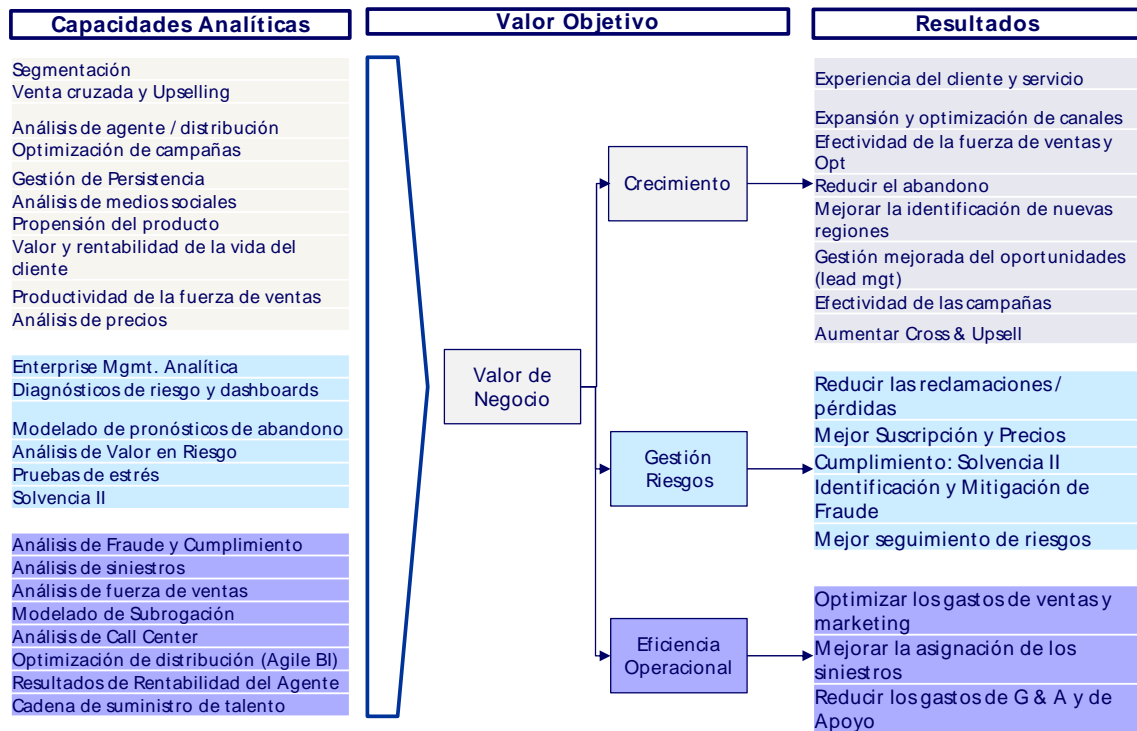
Existen usos generalistas agnósticos al sector/industria tales como enfocarnos a una mejora del crecimiento mediante predictiva de propensión al abandono que nos trata, o al de la compra para realizar venta cruzada. En caso de orientarnos a una mayor eficiencia en la gestión de riesgos, podemos aplicar estos principios a mejorar los ratios de identificación de fraude, mejora en la suscripción del riesgo y/o definición del precio.

En el caso de perseguir el objetivo de eficiencia interna, se pueden considerar casos de uso tales como la mejora en la asignación de siniestros a distin-

tos agentes, o proveedores a una solicitud de asistencia según distintos ratios y principios definidos.

El siguiente diagrama muestra un marco de descripción de potenciales capacidades analíticas, que pueden ser generalistas o específicas para entidades aseguradoras y financieras así como, dependiendo del valor objetivo que cada entidad busque, podrá identificar resultados en distintos ámbitos:

Gráfico 4. Capacidades de analítica avanzada y potenciales beneficios



Fuente: Propia

## 4.2 Análisis del estado de madurez de analítica avanzada en el sector asegurador

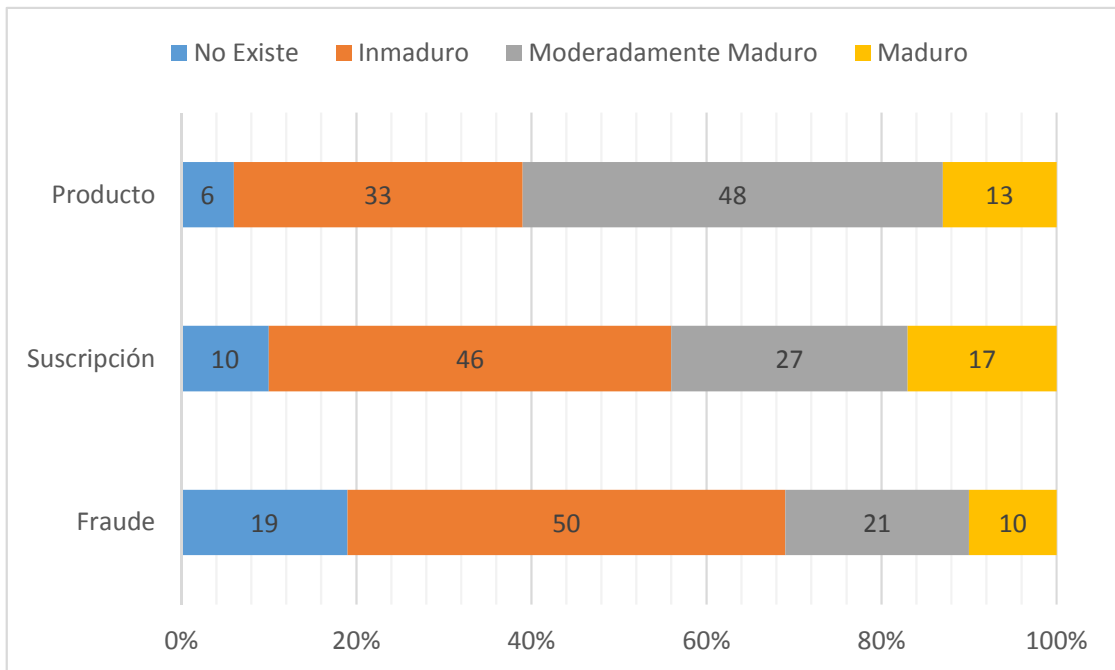
En esta sección hacemos un barrido a los casos de uso más comunes del sector. Existe un rápido crecimiento en el uso estratégico de los datos en toda la industria y estas están aprovechando nuevas fuentes de datos (p.ej. datos sociales y públicos, y datos de nuevas fuentes como internet de las cosas), que ayudarán a impulsar la optimización del negocio y ampliar el uso de analítica a nuevos casos de uso.

En una encuesta que publica la prestigiosa empresa de investigación y consultoría Gartner fechada recientemente, se ilustra el interés por la inteligencia empresarial (BI) y la analítica avanzada, que aumentan constantemente año tras año. Casi la mitad de los encuestados (47%) informó estas iniciativas se encontraban entre las tres áreas tecnológicas más importantes para la nueva inversión, muy por delante de los competidores más cercanos tales como digitalización y marketing digital

(27%) y cloud computing (23%). La encuesta también mostró que el 79% de los encuestados consideró que la analítica avanzada tenía el potencial de cambiar sus organizaciones durante los próximos cinco años. Ninguna otra tecnología tiene semejante expectativa de influenciar un cambio en la industria en esta magnitud.

El diagrama a continuación muestra las tres áreas más representativas del sector tras una encuesta realizada a un amplio espectro de aseguradoras, así como la madurez que comunicaron cada una de las encuestados.

Gráfico 5: Gartner Data and Analytics Summit 2017 Encuesta Resultado: Advanced Analytics Maturity in Insurance



Fuente: Gartner (Junio 2017)

#### 4.2.1 Analítica en el ámbito de definición de producto

El caso más representativo según el estudio del apartado anterior es el análisis avanzado para la gestión inteligente de la elaboración de nuevos productos. A medida que maduran las estrategias digitales de las aseguradoras, estas se están enfocando cada vez más en la innovación de productos, lo que incluye traer al mercado propuestas que satisfagan las necesidades de segmentos de clientes únicos, como los millennials y los consumidores digitales. Que soporten las necesidades no satisfechas (p. ej. seguros a la carta) y que se basen en el pago por uso.

Con esto, las entidades deben asegurarse que las nuevas propuestas son rentables y que los nuevos productos no erosionen la rentabilidad de los productos antiguos. También deben poder evaluar el valor de los productos existentes para que sepan cuándo retirarlos del mercado.

¿Qué conlleva esto desde el punto de vista técnico y desde el prisma de la función que desempeña el autor (sistemas de información) ? Pues poder desarrollar capacidades que den visibilidad en tiempo real del rendimiento del producto y predictivas para probar introducciones de nuevos productos. Mejorar las plataformas existentes para apoyar la gestión del ciclo de vida del producto, incluyendo análisis avanzados de productos que provean análisis del volumen de ventas, análisis del ROI entre otros. En el caso de lanzamiento de nuevos productos poder realizar análisis de impacto producto-producto, rentabilidad, proyección y modelación de la rentabilidad de los precios.

El objetivo es el de proveer de plataformas de análisis que se integren estrechamente con los sistemas de taller de producto así como los sistemas de administración de pólizas convencionales.

#### **4.2.2 Analítica en la suscripción**

Mientras que la rentabilidad de la suscripción ha sido un importante foco de las aseguradoras desde hace ya algún tiempo, la aplicación de nuevos tipos y orígenes de datos junto con el uso de analítica avanzada es aún bastante incipiente en el sector.

Los departamentos de suscripción continúan buscando nuevas formas para complementar las prácticas tradicionales y poder proveer decisiones precisas y automatizadas.

El objetivo en este ámbito de aplicación es el de proveer de motores de decisión que puedan integrarse en las herramientas de los actuarios y suscriptores y utilizar la analítica avanzada para ayudar en el análisis e interpretación de los datos. Este tipo de motor de decisión puede utilizar tanto Big Data como inteligencia artificial y la combinación de ambos para entregar más conocimiento del cliente, tanto a los sistemas automatizados de suscripción como a los colaboradores que deban tomar decisiones.

#### **4.2.3 Analítica en la detección de fraude**

Las tasas de fraude dirigido a las aseguradoras conllevan consigo volúmenes de gasto que, en un sector con márgenes cada vez más ajustados, tienen una influencia importante en la cuenta de resultados. Se prevee que el nivel de fraude pueda ir en aumento debido al incremento del volumen de negocio que se realiza electrónicamente y la cantidad de riesgos emergentes como resultado de transacciones cada vez más digitales.

El análisis de fraude tradicional, que utilizan patrones históricos, procesamiento por lotes y conjuntos de datos limitados ya no son adecuados. No ofrecen la precisión necesaria para el mercado digital de hoy, y conllevan demasiado tiempo para identificar patrones de fraude.



Mediante nuevos procedimientos analíticos y tecnologías para analizar una gama más amplia de datos en tiempo real, y la utilización de modelado predictivo y *machine learning* para la detección de anomalías (en términos estadísticos, outliers) se consigue una capacidad de acierto más elevada, y un menor número de falsos positivos.

En definitiva, se provee de eficiencia a los equipos de investigación de fraude a analizar aquellos casos que tienen mayor probabilidad de ser identificados como fraude y a obviar aquellos que pudieran ser falsos positivos, repercutiendo en un menor impacto a las arcas de la entidad y una mejor experiencia hacia el cliente.

En la práctica, esto requiere integrar los sistemas predictivos con aquellos sistemas de gestión de siniestros, en su gran mayoría sistemas de gestión de flujos de trabajo (Business Process Management o WorkFlows) para que tanto los tramitadores de siniestros como los gestores de fraude puedan conocer, en tiempo real tras la apertura de un nuevo siniestro, cuál es la probabilidad que este se trate de un fraude y cómo deben proceder con el mismo de la manera más eficiente posible.

### **4.3 Otros ejemplos prácticos**

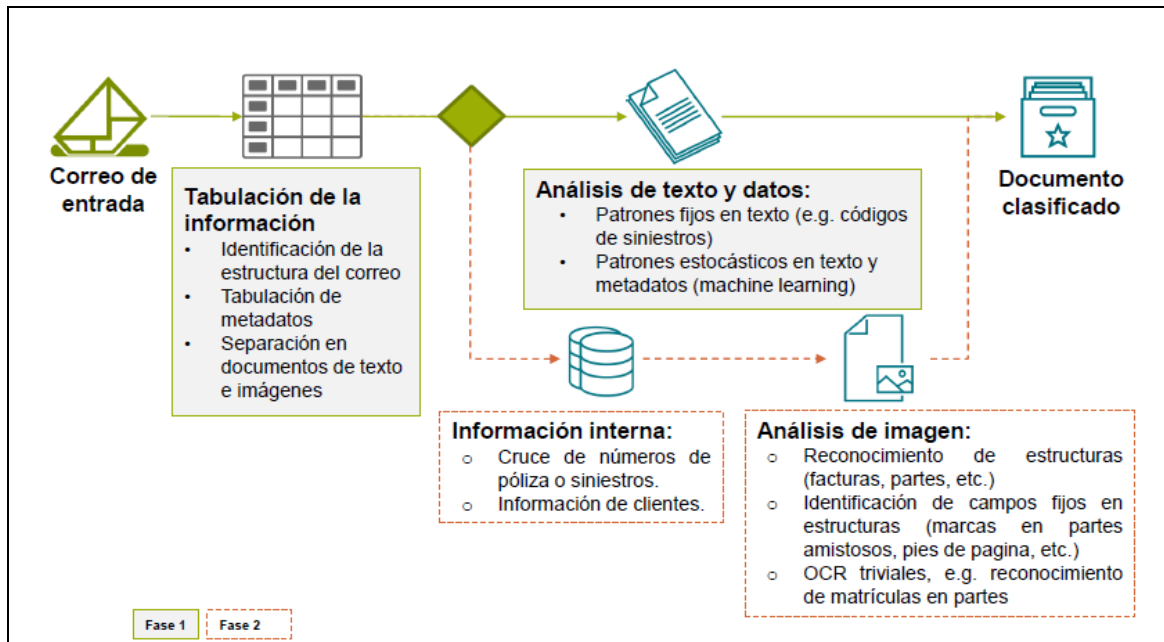
Esta sección describe en un nivel de detalle algo mayor dos casos prácticos realizados recientemente para proveer al lector en más extensión principios, objetivos y resultados empíricos

#### **4.3.2 Indexación documental**

Mientras que el ritmo de digitalización de la industria aseguradora es acelerado, existe aún multitud de información y documentación no estructurada que se recibe anualmente en las entidades. Documentos físicos, correos electrónicos, documentos pdf y otros se reciben día a día con múltiples propósitos y estructuras que requieren, en su casi totalidad, de la digestión, clasificación/categorización y almacenamiento previo a la provisión de un servicio. Partes de accidente de automóvil, solicitudes de información, de cancelación, informes periciales, facturas de proveedores de asistencia, atención a un siniestro u otros. Esto en mayoría de los casos, con un esfuerzo muy intensivo de mano de obra.

Existen ya aplicaciones de modelos analíticos basados en tecnología de big data que, junto modelos analíticos avanzados permiten alcanzar cotas de clasificación documental con una precisión superior al 90% y cometiendo un 2% de errores, lo que se estima 7,5 veces menos que los que realizarían un humano. Esto se ha conseguido para la ingestión de un estado inicial que gestiona un 30% de la ingestión de una aseguradora internacional con amplia presencia mundial. Los beneficios son claros, reducción de estructura, mayor velocidad en la ingestión y reducir riesgo operativo de una incorrecta indexación documental.

Gráfico 6. Enfoque del proceso de fuentes de información desestructurada



Fuente: Propia

### 4.3.3 Scoring de impago

Un segundo ejemplo práctico utilizado dentro del sector asegurador es la utilización de modelos para valorar cuál es la probabilidad de impago de un cliente potencial con el objetivo de reducir el ratio combinado de su cartera del ramo de autos con pagos de prima fraccionada.

El modelo predictivo tiene como propósito asignar una puntuación de impago a cada cliente en el momento en que contrata una póliza de seguro de automóvil para poder tomar la decisión desde la entidad.

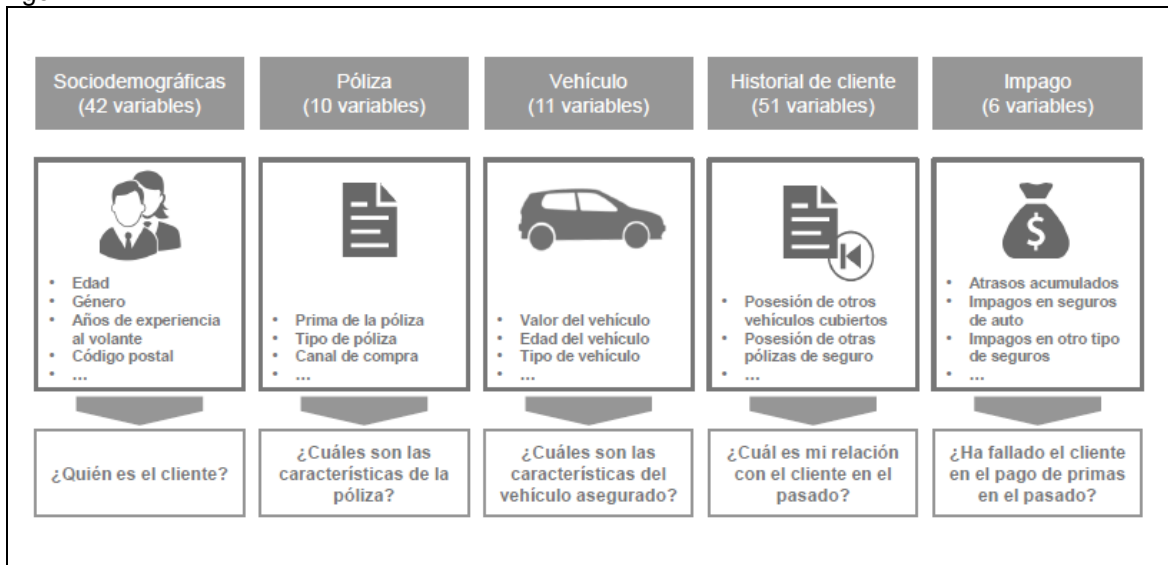
Los beneficios son diversos, desde mejoras del conocimiento del cliente que pueden ser utilizadas en iniciativas operativas y de marketing sin necesidad de modelos analíticos, así como definición de políticas de fraccionamiento de pago a nivel micro-segmento específico de riesgo evitando la anti selección.

En este caso se utilizó información interna (existente en la entidad en referencia al cliente y a los productos a asegurar) así como información pública externa tales como:

- Características socio-demográficas provenientes del INE tales como censo de población y vivienda, con granularidad a nivel de sección censal
- Tasas de desempleo del área de vivienda del cliente como los provistos por los municipios
- Niveles medios de ingresos provistos a nivel de municipio

Durante la fase de preparación modelación se definió una serie de variables de diferentes tipos que se resume en el siguiente gráfico.

Gráfico 7. Orientación sobre variables consideradas en la modelización del *scoring* de impago



Fuente: Propia

De estos, las variables que se identificaron como más relevantes para la propensión de impago resultaron ser en este orden, la edad del vehículo, el historial de accidentes así como el tipo de póliza a contratar. De modo que vehículos con más de 15 años, clientes con un mal historial de siniestros o las pólizas a terceros eran los grupos de mayor riesgo de impago.

En referencia a la relevancia de los factores externos, municipios con tasa de desempleo entre 20% al 55% mostraron un mayor sensibilidad mientras variables como el volumen medio de ingresos demostró no tener influencia significativa.



## 5. Modelo de propensión al abandono de auto

Como comentábamos en el capítulo anterior, el objetivo de esta tesis es el de una vez contextualizado el ámbito técnico y los distintos y múltiples casos de uso, adentrarnos un poco más en un caso concreto para conocer la idiosincrasia de la elaboración de un modelo avanzado.

La cantidad de detalle e información que puedo redactar en esta tesis que se publica con posterioridad es limitada, ya que el capital intelectual de estos modelos no puede compartirse ni publicarse sin revelar lo que puede ser una ventaja competitiva o diferencial. Lo que sí que detallaremos es información parcial y general que cumpla con el requisito de conocer mas detalladamente tanto el proceso de elaboración como la información relativa al caso, sin revelar o permitir la duplicación del mismo.

Existe un proceso de elaboración / modelización que se compone de seis fases principales comunmente utilizadas que son agnósticas al caso de uso que se elabore. Estas son, a grandes rasgos: el entendimiento de negocio, el entendimiento de los datos, la preparación de los datos, la modelización, la evaluación y por último el despliegue. Cada una de estas fases se descompone de distintos actividades que detallamos a continuación junto con la definición para el caso de propensión al abandono.

### 5.1 El entendimiento de negocio

#### 5.1.1 Definición del objetivo y criterio de éxito

En primera instancia debemos definir cuál es el objetivo y los criterios por los que se considerará que el modelo será acertado/exitoso.

**Objetivo:** Poder anticipar en un periodo prudencialmente anterior a la renovación (se considera dos meses) de una póliza de automóvil, cuál es la probabilidad que nuestro cliente desee abandonar la entidad. En otras palabras, la tasa de cancelación de clientes (en inglés *churn rate* o *attrition rate*).

**Criterio de Éxito:**

- Utilizar la probabilidad de *churn* en una campaña, junto con la inclusión de esta medición en el DataLake para su uso con otros fines comerciales.
- Construir un indicador para evaluar el modelo. Básicamente el grado de acierto del modelo.

#### 5.1.2 Evaluación de la situación y aproximación

Una vez definidos los objetivos, se realiza una evaluación de la situación, de modo que se identifica qué compone el inventario de información, las limita-

ciones que pueda tener, así como cuál será la aproximación para la definición del modelo.

En nuestro caso disponemos de un DataLake, que no es otra cosa que un repositorio de datos preparado para la explotación masiva, con información extensa proveniente de los distintos sistemas transaccionales. En el apartado de las limitaciones nos encontramos con algunas tales como identificar si se dispone de toda la información requerida. Situaciones como la disponibilidad de información de los distintos segmentos como pudiera ser vida y no-vida, si se requiere filtrar la información para obviar la que se disponga de distintas entidades, si existe información parcial en alguno de los ámbitos por razones históricas, son posibles limitaciones que se identificarán durante estos primeros pasos.

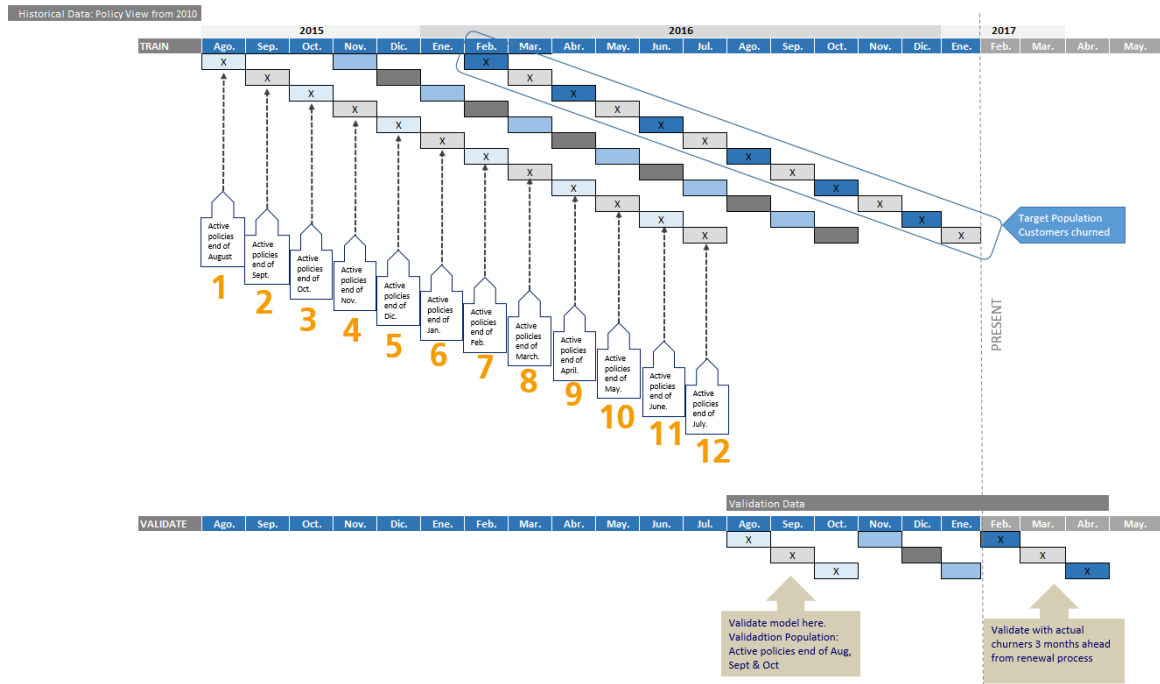
### **5.1.3 Definición de los objetivos de minería de datos, y descripción del alcance**

Esto es la definición de las variables iniciales a considerar, el horizonte de tiempo a analizar, cuáles serán las herramientas y las técnicas a utilizar entre otros. En el caso que nos ocupa se utilizará un universo de pólizas de 12 meses. Pólizas en vigor en agosto de 2015 que debieran renovar desde Noviembre 2015 hasta Octubre de 2016. Es decir, para cada póliza se utiliza el histórico de un año aproximadamente.

El modelo debe identificar la proporción de clientes del pasado que abandonaron la póliza P y son similares a mi cliente X actual. El cliente X, ¿cómo se comportará en el futuro? ¿Contratará o cancelará? Esto conlleva la definición de una variable objetivo, que simboliza que el cliente anula (1) o sigue en vigor (0). La dificultad radica en calcular correctamente esta variable.

La propensión se calcula mediante las observaciones reales previas. Como hemos observado qué hicieron en el pasado clientes similares a nuestro cliente X, sabemos cuánto terminaron cancelando. Esta proporción es nuestra predicción de propensión. Usamos esta proporción de cancelación del pasado para predecir la propensión en el futuro

De manera gráfica podemos ilustrarlo como sigue:



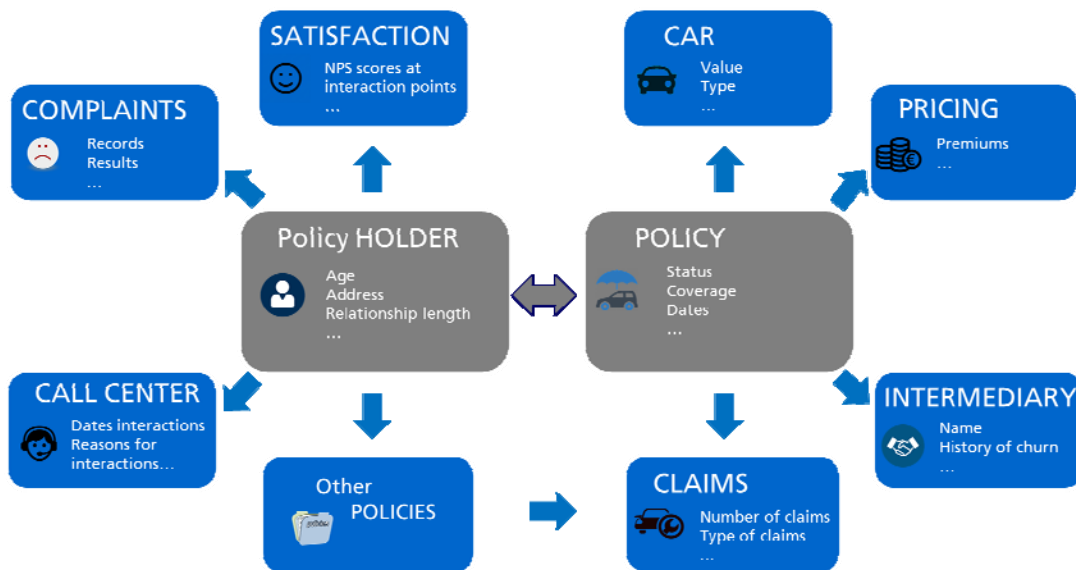
Las herramientas utilizadas son software libre estándar de mercado, en concreto la utilización de R-Studio para la visualización, gestión de calidad del software e informes, y la utilización de Python para el procesamiento de datos, construcción de modelos y su posterior validación.

## 5.2 Entendimiento y preparación de la información/datos

Tras la fase anterior cabe definir cuál es la información que requeriremos para la elaboración del modelo predictivo así como la definición de la variable objetivo.

Estas fases aglutinan el mayor esfuerzo en la identificación y depuración de los datos que obviaremos en esta sección pero que tienen como objetivo el de identificar un set de datos final con el que se procederá a la siguiente fase de modelización. El data set que se utilizó empíricamente para la elaboración de este incluyó centenares de variables de distinto índole (más de 260), agrupadas en lo que se denomina en el argot técnico como botellas, que no es otra cosa que una agrupación lógica de datos.

A modo de ejemplo, se detalla una definición parcial de qué datos se utilizaron para este propósito:



Fuente: Zurich España

**Información y variables relativas al ramo de automóvil:** Esto es, toda la información relativa al cliente tal como información socioeconómica (dirección, provincia, edad, estado civil) y toda extensión factible para este tanto en disposición de la entidad como en organismos de información pública.

**Información sobre el bien asegurado:** En el caso que nos ocupa, toda la información disponible sobre el automóvil tal como la marca, modelo, potencia, antigüedad y otros. Sin diferenciar en exceso a la información utilizada para el cálculo del riesgo y la prima utilizado en los departamentos actuariales de la entidad.

**Intermediación:** Información relativa al canal de comercialización de la póliza en su máxima extensión. Si esta se obtuvo mediante mediación agencial, afiliados, brokers u otros. Así como su localización, histórico e información ampliada.

**Información relativa a la Póliza:** Información de la póliza per sé, incluyendo desde garantías y coberturas como métodos de pago, antigüedad de la primera emisión, incidentes u ocurrencias de pago, entre otros.

**Información relativa al historial de siniestros del cliente:** Número de siniestros, culpabilidad o no de los mismos, fecha de ocurrencia y estado de los mismos, así como otra información complementaria en este ámbito tanto propia como provista por terceros.

**Prima:** Información relativa a la prima, prima histórica, actual y a renovación. Así como información comparativa de la misma hacia mercado.

**Portfolio de productos:** Información histórica sobre pólizas del ramo de automóvil u otros, tanto activos como inactivos. Incluyendo información relativa a otros segmentos si procede.



**Actividad del cliente:** Actividad y transaccionalidad del cliente con la compañía, en distintos canales y propósitos.

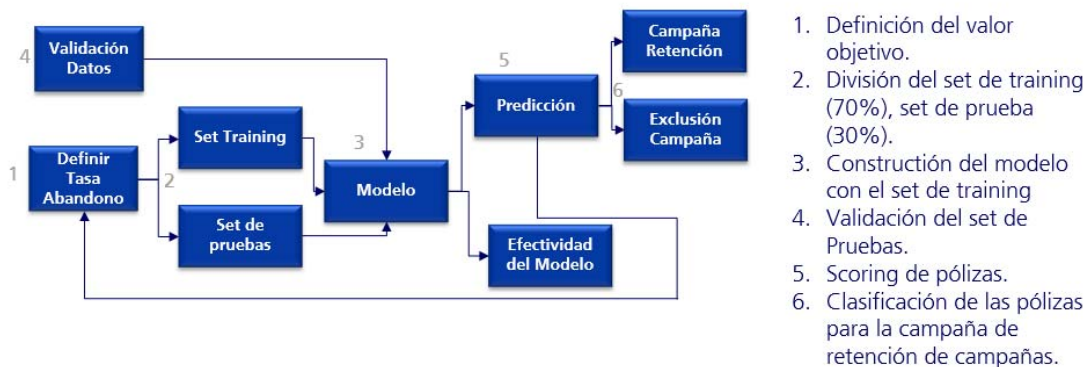
**Satisfacción del cliente:** Información relativa (si existe) sobre el *Net Promoting Score*, cuya valoración indica el grado de satisfacción del cliente con la compañía.

### 5.3 Modelización y Evaluación

El proceso de construcción del modelo, así como la validación también acostumbra a responder un mismo proceso. A partir de la definición del objetivo, la información histórica disponible se acostumbra a segmentar. Una parte se utiliza para la construcción (en el caso que nos ocupa un 70%) y se aparta una segunda parte que será con la que, una vez definido el modelo, se utilizará para realizar las pruebas y validar la tasa de acierto.

En nuestro caso el resultado de la predicción se utilizará para discernir la lógica a aplicar en las campañas de retención de clientes.

El diagrama a continuación ilustra el proceso utilizado y la aplicación ejemplar de utilización:



La construcción del modelo en este caso se realizó mediante la utilización de algoritmos avanzados tales como *random forest*, *boosting trees* y redes neuronales, entre otros cuyo resultado siempre debiera evocar un score / probabilidad de abandono.

Con posterioridad se utiliza la probabilidad de cancelación observada en el pasado para cada segmento y comprobamos qué proporción de clientes de cada segmento cancelaron el productos con posterioridad a la fecha de referencia.

Esta proporción será la que utilicemos como predicción de la propensión de cancelación/abandono para clientes actuales

## 5.4 Despliegue

Un apartado que merece un capítulo a parte es el despliegue de los modelos. El proceso que hemos explicado en este capítulo concluye en un modelo que, dados unos datos de entrada, concluye unos resultados que proveen de una inteligencia que aplicada a la empresa nos pueda dar una ventaja competitiva.

Pero estos modelos no son estáticos, se mejoran progresivamente, se ajustan según los resultados. Algunos de manera automática o asistida, y otros mediante aplicando esfuerzos en incremento o depuración de datos, nuevas variables, versionados u otros. Por tanto, al igual que los sistemas informáticos convencionales, estos requerirán todo un marco de configuración y gestión de versiones, código de extracción de información, transformación y carga de datos (ETL, siglas en inglés de Extraction Transformatio and Load), así como arquitecturas que permitan la ejecución online de los modelos a petición.

Sin querer entrar en tecnicismos, del mismo modo que solicitamos a nuestros sistemas una cotización para un póliza de automóvil o una emisión en segundos, debemos de ser capaces también de, a partir de un cliente, solicitar la ejecución de este modelo online para identificar su propensión al abandono.

Esto se podrá integrar en tantos procesos como decidamos. Inicialmente comentábamos que podrían utilizarse para las campañas de retención pero, tendría sentido utilizarlo como una variable más a la hora de calcular la prima de renovación ?.

El despliegue objetivo de modelos de analítica avanzada debe ser aquél que, utilice la información masiva pero que provea de un marco donde se puedan consumir los servicios de manera inmediata y dentro de un marco que soporte la actualización y gestión de mantenimiento y mejora de los modelos.

## 6. Organización

Un ámbito importante a tratar cuando una empresa decide acometer la explotación de analítica avanzada de grandes volúmenes de datos es cómo afrontarlo desde el punto de vista de la organización. Existen tantos modelos como uno desee, desde la externalización parcial de los servicios, colaboraciones con entidades especializadas o instituciones educativas hasta la total internalización de los mismos. En este sentido, no tiene diferenciación en cómo uno pueda definir en la organización de otros servicios.

Lo que sí quizás sea un factor común independiente a la organización y pueda resultar de interés es la composición de tres ámbitos de conocimiento diferenciado. En la experiencia del autor, la industrialización de un servicio se compone en tres funciones/equipos diferenciados:

- **Tecnologías de la información:** Se requiere un equipo con amplios conocimientos en tecnologías de la información. Por un lado requiere el conocimiento de la información y localización de los datos requeridos, para poder extraer y tratar esta información a distintas tecnologías de almacenamiento de datos que variarán según cuál sea el modelo a realizar. Esto incluye conocimientos desde bases de datos relacionales convencionales a arquitecturas y estándares de almacenamiento no estructurados de datos. Conocimientos en extracción, tratamiento y carga de información, depuración de datos entre otros. Será en cierta medida, quién aporte los datos para su modelización posterior.
- **Analítica avanzada:** En el desarrollo de los distintos modelos, ya sean predictivos, descriptivos, diagnósticos o de otra índole, se identifica la necesidad de disponer de perfiles expertos en modelización y analítica, que confluye conocimientos matemáticos y estadísticos. Son los especialistas en comprender la idiosincrasia de cada caso de negocio, los datos de los que se sustentan, la modelización mediante distintas técnicas así como la evaluación de cada iteración hasta la definición y posterior mantenimiento de los modelos puestos en marcha. Este perfil se ha generalizado recientemente con el nombre de *Data Scientist* o *Data Analyst* y debe iniciar en cada caso por el entendimiento del caso, la interpretación de los datos y la información, la preparación de la información más allá de la provista por los equipos de sistemas, así como la evaluación de los modelos, aprendizaje de los mismos hasta su puesta en producción.
- **Profesionales del ámbito de aplicación:** El tercer actor es el receptor del servicio, que puede diferir según cada caso y que debe transmitir al equipo anterior todo el conocimiento sobre su operación y sobre la necesidad. En caso de requerir un modelo de identificación de potencial fraude, será el equipo de gestión de siniestros y el equipo de gestión de fraude quienes deberán proveer cuál es la necesidad, la información disponible, la deseable, los impedimentos y todo tipo de in-

formación que permita a los Data Scientists entender la idiosincrasia y que ahí parta la metodología de desarrollo. En el caso de tratarse de un modelo de propensión a la fuga de clientes será el departamento de ventas y distribución junto con marketing quienes transmitirán esa necesidad y conocimiento. Este tercer actor será quien, mediante la automatización de modelos de analítica avanzada apalancada en volúmenes extensos de información se beneficie de la inteligencia avanzada automatizada.

La identificación del problema puede provenir de distintos ámbitos de la organización pero su desarrollo siempre pivota entre estas tres funciones anteriormente descritas.

## **7. Amenazas al BIG DATA: Reglamento General de Protección de Datos (GDPR)**

Como en la mayoría de iniciativas apalancadas en avances tecnológicos aparecen temas relacionados con el impacto que estos pueden tener tanto en el ámbito empresarial como en el social. Y en eso Big Data no es distinto.

Si bien es cierto que la explosión de la generación de los datos ha sido muy abrupta en los últimos años, también lo ha sido la acesibilidad y almacenamiento de información relativa a personas privadas. Información no sólo sobre la identificación de un individuo (léase nombre, dirección, identificación fiscal/nacional) sinó sobre sus hábitos tales como geolocalización, hábitos de consumo mediante explotación de información de crédito, historiales de navegación, u otros. Puede llegar a ser abrumadora la cantidad de información que se puede llegar a conocer de nosotros a través del rastro digital que realizamos, equivalente a la huella que dejamos a cada paso de damos.

Muchos han sido los casos en los que se han filtrado u utilizado de manera irregular datos sobre personas físicas con fines no deseados, o se han publicado estos de manera ilícita. La sensibilidad sobre la privacidad varía de manera ostensible de individuo a individuo.

Para regular y proteger la información de las personas físicas se elaboró la ley en el estado Español denominada Ley Orgánica de Protección de Datos (LOPD) de Caracter Personal fechada en 1999 y que entró en vigor a principios del 2000, derogando la ley que hasta ese momento regulaba en España el tratamiento automatizado de datos personales (conocida como lortad). Con la entrada en vigor de la LOPD se cumplía con la Directiva Europea para todos los estados miembros publicada en 1995.

Pues bien, la nueva Directiva Europea conocida como GDPR (Siglas en inglés de General Data Protection Regulation) extiende aún más a su predecesora para proteger aún más a todos los ciudadanos de la UE contra la privacidad y las violaciones de datos en un contexto cada vez más orientado a los datos. Sustituye la directiva Europea que se estableció en 1995 y ha sido diseñada también para armonizar las leyes de privacidad de datos en toda Europa.

Entrará en vigor 20 días después de su publicación en el Diario Oficial de la UE y se aplicará directamente en todos los Estados miembros dos años después de esta fecha. Esto es el próximo 25 de mayo de 2018.

Uno de los puntos que varían sustancialmente es el régimen sancionador, que se vuelve significativamente más severo, sobre todo para aquellas empresas que tenga mayor facturación, y que afectarán tanto a los responsables como a los encargados de protección de los datos. Las multas podrán alcanzar hasta los 20 millones de euros o el 4% del volumen de negocios total anual del ejercicio financiero anterior.

Aunque los principios clave de la privacidad de los datos siguen siendo válidos para la directiva anterior, se han propuesto muchos cambios a las políticas reguladoras. Nuevos elementos como el derecho al olvido y el derecho a la portabilidad, que mejoran la capacidad de decisión y control de los ciudadanos sobre los datos personales que confían a terceros.

Uno de los aspectos esenciales del reglamento es que se basa en la prevención por parte de las organizaciones que tratan datos, la conocida como responsabilidad activa. "Las empresas deben adoptar medidas que aseguren razonablemente que están en condiciones de cumplir con los principios, derechos y garantías que el GDPR establece".

Los puntos claves así como la información sobre los impactos que tendrá sobre el negocio son amplios y serían susceptibles de una tesis amplia pero para el propósito que nos ocupa en esta ocasión, realizo una síntesis de las regulaciones que pueden limitar de una manera substancial o que requerirán una adecuación de los procesos y servicios actuales el ámbito de Big Data. Se resumen en los siguientes ámbitos:

**Alcance geográfico:** El principal cambio en este ámbito radica en que el reglamento se aplica no solo a las organizaciones establecidas en la comunidad Europea, sino también a las organizaciones establecidas fuera de la Unión Europea si recogen o procesan datos personales de residentes en la UE. Es decir, hasta ahora, si una empresa (p.e. Google) operaba en Europa, podía no cumplir las directrices Europeas al tener sus matrices fuera de la Unión.

**Consentimiento:** Este es un apartado fundamental, y que establece que los individuos deben conceder consentimiento explícito (remarco el concepto de explícito) para los datos recopilados y también para los fines en que se utilicen los datos (artículo 7, definidos en el artículo 4). Los controladores de datos (figura que añade el reglamento) deben ser capaces de probar el "consentimiento" (opt-in) y el consentimiento puede ser retirado. Es decir, con la entrada de la nueva ley, el ciudadano tiene que dar consentimiento y este no puede ser genérico como el que estamos ahora acostumbrados. El consentimiento no podrá ser dado por defecto y para fines genéricos sino que las entidades estarán obligadas a expresar el uso y detallarlo para que la conformidad sea clara y explícita siempre y cuando este no sea fundamentalmente requerido para proveer los servicios para los cuales se le contrata.

**Perfilado de datos:** Una selección del artículo 60 cita textualmente: "Los principios de tratamiento leal y transparente exigen que se informe al interesado de la existencia de la operación de tratamiento y sus fines. El responsable del tratamiento debe facilitar al interesado cuanto información complementaria sea necesaria para garantizar un tratamiento leal y transparente, habida cuenta de las circunstancias y del contexto específico en que se traten los datos personales. Se debe además informar al interesado de la existencia **de la elaboración de**

**perfiles y de las consecuencias de dicha elaboración.** Si los datos personales se obtienen de los interesados, también se les debe informar de si están obligados a facilitarlos y de las consecuencias en caso de que no lo hicieran. Dicha información puede transmitirse en combinación con unos iconos normalizados que ofrezcan, de forma fácilmente visible, inteligible y claramente legible, una adecuada visión de conjunto del tratamiento previsto. Los iconos que se presentan en formato electrónico deben ser legibles mecánicamente.

De suma importancia el extracto siguiente:

*(...) Todo interesado debe, por tanto, tener el derecho a conocer y a que se le comuniquen, en particular, los fines para los que se tratan los datos personales, su plazo de tratamiento, sus destinatarios, la lógica implícita en todo tratamiento automático de datos personales y, **por lo menos cuando se base en la elaboración de perfiles, las consecuencias de dicho tratamiento.** (..)*

Esto implica que a partir de la entrada en vigor de esta ley, los individuos deberán dar conformidad de manera explícita y clara cada uno de los propósitos de la utilización de los datos más allá del propósito/servicio que la empresa rinda. En el ámbito del sector asegurador, esto puede implicar que se deba solicitar conformidad para que su información sea utilizada de tanto detallada como anonimizada (léase perfilada) para la elaboración de modelos.

Por tanto y para concluir, dependiendo de cómo se acabe aplicando la ley en la práctica, la potencialidad de casos de uso que en la mayoría de los casos se realizan tanto con datos completos como por perfilados de datos podrían limitarse para proteger la privacidad del individuo.

¿Estará permitida la realización de modelos predictivos que perfilen la información para poder predecir la propensión al abandono? ¿Permitirán mis clientes que les solicite cada uno de los modelos y propósitos más allá de lo que requiere la empresa para desarrollar el servicio contratado? ¿Qué deberé realizar con los clientes existentes? ¿Debemos solicitar aceptación explícita activa previa?.

Aquí se encuentra la amenaza y como el Big Data puede verse merchado por la incapacidad de acceder al universo de datos de los que disponemos actualmente, o de la beligerancia con la que aceptamos implícitamente este tipo de procesamiento en la actualidad.





## 8. Capítulo de conclusiones

Espero que el lector pueda anticipar cuáles son las conclusiones habiéndolo leído la redacción hasta estas líneas. Las capacidades informáticas de la gestión de grandes volúmenes de información han evolucionado, las técnicas y capacidades de las distintas ciencias (matemáticas, estadísticas, actuariales, y tantas otras) están ya vinculadas a esta práctica y ya tenemos presentes en un amplio tejido empresarial departamentos de análisis avanzado de información, que tiene como cometido el elaborar análisis de distinto indole más allá del análisis descriptivo al que estábamos acostumbrados hasta la fecha.

Hay quienes auguran que el valor de una empresa estará vinculado al valor de sus modelos algorítmicos inteligentes, lo que bautizan como *algorithmic businesses*.

El sector asegurador debiera estar aventajado en este ámbito, porque la ciencia actuarial ya conjuga la aplicación de modelos matemáticos y estadísticos ligadas a la identificación y predicción de la materialización de un riesgo cumpliendo la teoría de los grandes números para poder dar pie a la definición de un precio que perpetúe la existencia de la industria. Ya disponemos en nuestras organizaciones perfiles que conjugan un amplio espectro de esta ciencia.

En el caso del modelo de propensión al abandono, nos encontramos en un escenario más dentro del casi infinito universo de posibilidades que brinda el Big Data. Demostramos a raíz de una experiencia empírica empresarial, que mediante la información interna y externa disponible se permite anticipar con un índice de acierto elevado, cuál va a ser el comportamiento de nuestros clientes.

La aplicación de modelos similares para su utilización en la mercadotecnia es un peldaño dentro de la larga cadena de ámbitos de aplicación, sin olvidarnos de los usos sociales y macroeconómicos que estos pueden aportar a la sociedad.



## 9. Bibliografía

### Artículos:

OLIVER RALPH. "Insurance and the Big Data Technology Revolution" Financial Times (24 Febrero 2017)

LOPEZ GARCIA, DAVID. "Analysis of the possibilities of use of Big Data in organizations" . (18 de octubre de 2014).

ALEGRE, I., ARIÑO M.ACANELA M.A. "Cómo sacar partido del análisis de datos. BIG DATA, por dónde empezamos". Revista antiguos alumnos IESE (Abril-Junio 2017)

LULL, EDUARDO. "Big Data Analysis to transform insurance Industry" Financial Times (16 Mayo de 2016)

MARR, BERNARD. "How Big Data is changing Insurance forever". Forbes (16 de Diciembre 2016)

GUILLEM, M. "Big Data en Seguros". Revista Índice (Abril 2016)

### Libros:

CABALLERO R., MARTIN E. "Las bases de big data". Ed. La catarata, 2015

PÉREZ MÁRQUEZ M. "BIG DATA. Técnicas, herramientas y aplicaciones" Ed. Alfaomega, 2015.

### Fuentes de internet:

MATTHEW KOMOROWSKI. *A history of storage cost*

<<http://www.mkomo.com/cost-per-gigabyte>>

Fecha de la consulta: 28 de abril 2017

Hard Drive Storage Cost

<http://ns1758.ca/winch/winchest.html>

Fecha de la consulta: 20 de abril 2017

CHEESMAN P. Cambrian Explosion of Data

<[http://www.eetimes.com/author.asp?section\\_id=36&doc\\_id=1330462](http://www.eetimes.com/author.asp?section_id=36&doc_id=1330462)>

Fecha de la consulta: 24 de abril 2017

EXASTAX. Top 7 Data Use Cases in Insurance Industry.

<https://www.exastax.com/big-data/top-7-big-data-use-cases-in-insurance-industry/>

Fecha de la consulta: 4 de Junio 2017

**Fuentes Oficiales:**

LEY ORGÁNICA 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal («BOE» núm. 298, de 14 de diciembre de 1999, páginas 43088 a 43099 (12 págs.)

REGLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 27 de abril de 2016

## **Oscar Pallisa Gabriel**

Durante mi carrera he tenido diferentes roles relacionados con la gestión de tecnologías de la información. Fui nombrado director del departamento de tecnologías de información de Zurich General Insurance Group en España en 2014, donde desempeñé diferentes funciones de liderazgo en programas internacionales de transformación.

Antes de desempeñar el cargo en Zurich España, dirigí el desarrollo de la plataforma de gestión de seguros generales en Latinoamérica tras realizar un ejercicio similar para Zurich General Insurance para Suiza.

Anteriormente, dirigí en el Sector Farmacéutico, la creación de un Centro de Excelencia de prestación de servicios a Europa y Mercados Emergentes y anteriormente tuve diversas experiencias de desarrollo de sistemas de información en empresas de consultoría e integradores de sistemas en el Reino Unido, USA, Canadá, Australia, y distintos países Europeos.

Ingeniero en Sistemas de Información en la Universidad Politécnica de Cataluña de Barcelona, y BSC (Hons) en Informática en la Universidad South Bank de Londres y MBA de la Universidad Pompeu Fabra.



**COLECCIÓN “CUADERNOS DE DIRECCIÓN ASEGURADORA”**  
Máster en Dirección de Entidades Aseguradoras y Financieras  
Facultad de Economía y Empresa. Universidad de Barcelona

**PUBLICACIONES**

- 1.- Francisco Abián Rodríguez: “Modelo Global de un Servicio de Prestaciones Vida y su interrelación con Suscripción” 2005/2006
- 2.- Erika Johanna Aguilar Olaya: “Gobierno Corporativo en las Mutualidades de Seguros” 2005/2006
- 3.- Alex Aguyé Casademunt: “La Entidad Multicanal. Elementos clave para la implantación de la Estrategia Multicanal en una entidad aseguradora” 2009/2010
- 4.- José María Alonso-Rodríguez Piedra: “Creación de una plataforma de servicios de siniestros orientada al cliente” 2007/2008
- 5.- Jorge Alvez Jiménez: “innovación y excelencia en retención de clientes” 2009/2010
- 6.- Anna Aragonés Palom: “El Cuadro de Mando Integral en el Entorno de los seguros Multirriesgo” 2008/2009
- 7.- Maribel Avila Ostos: “La tele-suscripción de Riesgos en los Seguros de Vida” 2009/20010
- 8.- Mercé Bascompte Riquelme: “El Seguro de Hogar en España. Análisis y tendencias” 2005/2006
- 9.- Aurelio Beltrán Cortés: “Bancaseguros. Canal Estratégico de crecimiento del sector asegurador” 2010/2011
- 10.- Manuel Blanco Alpuente: “Delimitación temporal de cobertura en el seguro de responsabilidad civil. Las cláusulas claims made” 2008/2009
- 11.- Eduard Blanxart Raventós: “El Gobierno Corporativo y el Seguro D & O” 2004/2005
- 12.- Rubén Bouso López: “El Sector Industrial en España y su respuesta aseguradora: el Multirriesgo Industrial. Protección de la empresa frente a las grandes pérdidas patrimoniales” 2006/2007
- 13.- Kevin van den Boom: “El Mercado Reasegurador (Cedentes, Brokers y Reaseguradores). Nuevas Tendencias y Retos Futuros” 2008/2009
- 14.- Laia Bruno Sazatornil: “L’ètica i la rentabilitat en les companyies asseguradores. Proposta de codi deontològic” 2004/2005
- 15.- María Dolores Caldés Llopis: “Centro Integral de Operaciones Vida” 2007/2008
- 16.- Adolfo Calvo Llorca: “Instrumentos legales para el recobro en el marco del seguro de crédito” 2010/2011
- 17.- Ferran Camprubí Baiges: “La gestión de las inversiones en las entidades aseguradoras. Selección de inversiones” 2010/2011
- 18.- Joan Antoni Carbonell Aregall: “La Gestió Internacional de Sinistres d’Automòbil amb Resultat de Danys Materials” 2003-2004
- 19.- Susana Carmona Llevadot: “Viabilidad de la creación de un sistema de Obra Social en una entidad aseguradora” 2007/2008
- 20.- Sergi Casas del Alcazar: “El PLAN de Contingencias en la Empresa de Seguros” 2010/2011
- 21.- Francisco Javier Cortés Martínez: “Análisis Global del Seguro de Decesos” 2003-2004
- 22.- María Carmen Ceña Nogué: “El Seguro de Comunidades y su Gestión” 2009/2010
- 23.- Jordi Cots Paltor: “Control Interno. El auto-control en los Centros de Siniestros de Automóviles” 2007/2008
- 24.- Montserrat Cunillé Salgado: “Los riesgos operacionales en las Entidades Aseguradoras” 2003-2004
- 25.- Ricard Doménech Pagés: “La realidad 2.0. La percepción del cliente, más importante que nunca” 2010/2011
- 26.- Luis Domínguez Martínez: “Formas alternativas para la Cobertura de Riesgos” 2003-2004
- 27.- Marta Escudero Cutal: “Solvencia II. Aplicación práctica en una entidad de Vida” 2007/2008
- 28.- Salvador Esteve Casablanca: “La Dirección de Reaseguro. Manual de Reaseguro” 2005/2006

- 29.- Alvaro de Falguera Gaminde: "Plan Estratégico de una Correduría de Seguros Náuticos" 2004/2005
- 30.- Isabel M<sup>a</sup> Fernández García: "Nuevos aires para las Rentas Vitalicias" 2006/2007
- 31.- Eduard Fillet Catarina: "Contratación y Gestión de un Programa Internacional de Seguros" 2009/2010
- 32.- Pablo Follana Murcia: "Métodos de Valoración de una Compañía de Seguros. Modelos Financieros de Proyección y Valoración consistentes" 2004/2005
- 33.- Juan Fuentes Jassé: "El fraude en el seguro del Automóvil" 2007/2008
- 34.- Xavier Gabarró Navarro: ""El Seguro de Protección Jurídica. Una oportunidad de Negocio"" 2009/2010
- 35.- Josep María Galcerá Gombau: "La Responsabilidad Civil del Automóvil y el Daño Corporal. La gestión de siniestros. Adaptación a los cambios legislativos y propuestas de futuro" 2003-2004
- 36.- Luisa García Martínez: "El Carácter tuitivo de la LCS y los sistemas de Defensa del Asegurado. Perspectiva de un Operador de Banca Seguros" 2006/2007
- 37.- Fernando García Giralt: "Control de Gestión en las Entidades Aseguradoras" 2006/2007
- 38.- Jordi García-Muret Ubis: "Dirección de la Sucursal. D. A. F. O." 2006/2007
- 39.- David Giménez Rodríguez: "El seguro de Crédito: Evolución y sus Canales de Distribución" 2008/2009
- 40.- Juan Antonio González Arriete: "Línea de Descuento Asegurada" 2007/2008
- 41.- Miquel Gotés Grau: "Assegurances Agràries a BancaSeguros. Potencial i Sistema de Comercialització" 2010/2011
- 42.- Jesús Gracia León: "Los Centros de Siniestros de Seguros Generales. De Centros Operativos a Centros Resolutivos. De la optimización de recursos a la calidad de servicio" 2006/2007
- 43.- José Antonio Guerra Díez: "Creación de unas Tablas de Mortalidad Dinámicas" 2007/2008
- 44.- Santiago Guerrero Caballero: "La politización de las pensiones en España" 2010/2011
- 45.- Francisco J. Herencia Conde: "El Seguro de Dependencia. Estudio comparativo a nivel internacional y posibilidades de desarrollo en España" 2006/2007
- 46.- Francisco Javier Herrera Ruiz: "Selección de riesgos en el seguro de Salud" 2009/2010
- 47.- Alicia Hoya Hernández: "Impacto del cambio climático en el reaseguro" 2008/2009
- 48.- Jordi Jiménez Baena: "Creación de una Red de Agentes Exclusivos" 2007/2008
- 49.- Oriol Jorba Cartoixà: "La oportunidad aseguradora en el sector de las energías renovables" 2008/2009
- 50.- Anna Juncá Puig: "Una nueva metodología de fidelización en el sector asegurador" 2003/2004
- 51.- Ignacio Lacalle Goría: "El artículo 38 Ley Contrato de Seguro en la Gestión de Siniestros. El procedimiento de peritos" 2004/2005
- 52.- M<sup>a</sup> Carmen Lara Ortiz: "Solventia II. Riesgo de ALM en Vida" 2003/2004
- 53.- Haydée Noemí Lara Téllez: "El nuevo sistema de Pensiones en México" 2004/2005
- 54.- Marta Leiva Costa: "La reforma de pensiones públicas y el impacto que esta modificación supone en la previsión social" 2010/2011
- 55.- Victoria León Rodríguez: "Problemàtica del aseguramiento de los Jóvenes en la política comercial de las aseguradoras" 2010/2011
- 56.- Pilar Lindín Soriano: "Gestión eficiente de pólizas colectivas de vida" 2003/2004
- 57.- Víctor Lombardero Guarner: "La Dirección Económico Financiera en el Sector Asegurador" 2010/2011
- 58.- Maite López Aladros: "Análisis de los Comercios en España. Composición, Evolución y Oportunidades de negocio para el mercado asegurador" 2008/2009
- 59.- Josep March Arranz: "Los Riesgos Personales de Autónomos y Trabajadores por cuenta propia. Una visión de la oferta aseguradora" 2005/2006



- 60.- Miquel Maresch Camprubí: "Necesidades de organización en las estructuras de distribución por mediadores" 2010/2011
- 61.- José Luis Marín de Alcaraz: "El seguro de impago de alquiler de viviendas" 2007/2008
- 62.- Miguel Ángel Martínez Boix: "Creatividad, innovación y tecnología en la empresa de seguros" 2005/2006
- 63.- Susana Martínez Corveira: "Propuesta de Reforma del Baremo de Autos" 2009/2010
- 64.- Inmaculada Martínez Lozano: "La Tributación en el mundo del seguro" 2008/2009
- 65.- Dolores Melero Montero: "Distribución en bancaseguros: Actuación en productos de empresas y gerencia de riesgos" 2008/2009
- 66.- Josep Mena Font: "La Internalización de la Empresa Española" 2009/2010
- 67.- Angela Milla Molina: "La Gestión de la Previsión Social Complementaria en las Compañías de Seguros. Hacia un nuevo modelo de Gestión" 2004/2005
- 68.- Montserrat Montull Rossón: "Control de entidades aseguradoras" 2004/2005
- 69.- Eugenio Morales González: "Oferta de licuación de patrimonio inmobiliario en España" 2007/2008
- 70.- Lluís Morales Navarro: "Plan de Marketing. División de Bancaseguros" 2003/2004
- 71.- Sonia Moya Fernández: "Creación de un seguro de vida. El éxito de su diseño" 2006/2007
- 72.- Rocio Moya Morón: "Creación y desarrollo de nuevos Modelos de Facturación Electrónica en el Seguro de Salud y ampliación de los modelos existentes" 2008/2009
- 73.- María Eugenia Mugerza Goya: "Bancaseguros. La comercialización de Productos de Seguros No Vida a través de redes bancarias" 2005/2006
- 74.- Ana Isabel Mullor Cabo: "Impacto del Envejecimiento en el Seguro" 2003/2004
- 75.- Estefanía Nicolás Ramos: "Programas Multinacionales de Seguros" 2003/2004
- 76.- Santiago de la Nogal Mesa: "Control interno en las Entidades Aseguradoras" 2005/2006
- 77.- Antonio Nolasco Gutiérrez: "Venta Cruzada. Mediación de Seguros de Riesgo en la Entidad Financiera" 2006/2007
- 78.- Francesc Ocaña Herrera: "Bonus-Malus en seguros de asistencia sanitaria" 2006/2007
- 79.- Antonio Olmos Francino: "El Cuadro de Mando Integral: Perspectiva Presente y Futura" 2004/2005
- 80.- Luis Palacios García: "El Contrato de Prestación de Servicios Logísticos y la Gerencia de Riesgos en Operadores Logísticos" 2004/2005
- 81.- Jaume Paris Martínez: "Segmento Discapacitados. Una oportunidad de Negocio" 2009/2010
- 82.- Martín Pascual San Martín: "El incremento de la Longevidad y sus efectos colaterales" 2004/2005
- 83.- Montserrat Pascual Villacampa: "Proceso de Tarificación en el Seguro del Automóvil. Una perspectiva técnica" 2005/2006
- 84.- Marco Antonio Payo Aguirre: "La Gerencia de Riesgos. Las Compañías Cautivas como alternativa y tendencia en el Risk Management" 2006/2007
- 85.- Patricia Pérez Julián: "Impacto de las nuevas tecnologías en el sector asegurador" 2008/2009
- 86.- María Felicidad Pérez Soro: "La atención telefónica como transmisora de imagen" 2009/2010
- 87.- Marco José Piccirillo: "Ley de Ordenación de la Edificación y Seguro. Garantía Decenal de Daños" 2006/2007
- 88.- Irene Plana Güell: "Sistemas d'Informació Geogràfica en el Sector Assegurador" 2010/2011
- 89.- Sonia Plaza López: "La Ley 15/1999 de Protección de Datos de carácter personal" 2003/2004
- 90.- Pere Pons Pena: "Identificación de Oportunidades comerciales en la Provincia de Tarragona" 2007/2008
- 91.- María Luisa Postigo Díaz: "La Responsabilidad Civil Empresarial por accidentes del trabajo. La Prevención de Riesgos Laborales, una asignatura pendiente" 2006/2007
- 92.- Jordi Pozo Tamarit: "Gerencia de Riesgos de Terminales Marítimas" 2003/2004
- 93.- Francesc Pujol Niñerola: "La Gerencia de Riesgos en los grupos multisectoriales" 2003-2004

- 94.- M<sup>ª</sup> del Carmen Puyol Rodríguez: "Recursos Humanos. Breve mirada en el sector de Seguros" 2003/2004
- 95.- Antonio Miguel Reina Vidal: "Sistema de Control Interno, Compañía de Vida. Bancaseguros" 2006/2007
- 96.- Marta Rodríguez Carreiras: "Internet en el Sector Asegurador" 2003/2004
- 97.- Juan Carlos Rodríguez García: "Seguro de Asistencia Sanitaria. Análisis del proceso de tramitación de Actos Médicos" 2004/2005
- 98.- Mónica Rodríguez Nogueiras: "La Cobertura de Riesgos Catastróficos en el Mundo y soluciones alternativas en el sector asegurador" 2005/2006
- 99.- Susana Roquet Palma: "Fusiones y Adquisiciones. La integración y su impacto cultural" 2008/2009
- 100.- Santiago Rovira Obradors: "El Servei d'Assegurances. Identificació de les variables clau" 2007/2008
- 101.- Carlos Ruano Espí: "Microseguro. Una oportunidad para todos" 2008/2009
- 102.- Mireia Rubio Cantisano: "El Comercio Electrónico en el sector asegurador" 2009/2010
- 103.- María Elena Ruiz Rodríguez: "Análisis del sistema español de Pensiones. Evolución hacia un modelo europeo de Pensiones único y viabilidad del mismo" 2005/2006
- 104.- Eduardo Ruiz-Cuevas García: "Fases y etapas en el desarrollo de un nuevo producto. El Taller de Productos" 2006/2007
- 105.- Pablo Martín Sáenz de la Pascua: "Solvencia II y Modelos de Solvencia en Latinoamérica. Sistemas de Seguros de Chile, México y Perú" 2005/2006
- 106.- Carlos Sala Farré: "Distribución de seguros. Pasado, presente y tendencias de futuro" 2008/2009
- 107.- Ana Isabel Salguero Matarín: "Quién es quién en el mundo del Plan de Pensiones de Empleo en España" 2006/2007
- 108.- Jorge Sánchez García: "El Riesgo Operacional en los Procesos de Fusión y Adquisición de Entidades Aseguradoras" 2006/2007
- 109.- María Angels Serral Floreta: "El lucro cesante derivado de los daños personales en un accidente de circulación" 2010/2011
- 110.- David Serrano Solano: "Metodología para planificar acciones comerciales mediante el análisis de su impacto en los resultados de una compañía aseguradora de No Vida" 2003/2004
- 111.- Jaume Siberta Durán: "Calidad. Obtención de la Normativa ISO 9000 en un centro de Atención Telefónica" 2003/2004
- 112.- María Jesús Suárez González: "Los Poolings Multinacionales" 2005/2006
- 113.- Miguel Torres Juan: "Los siniestros IBNR y el Seguro de Responsabilidad Civil" 2004/2005
- 114.- Carlos Travé Babiano: "Provisiones Técnicas en Solvencia II. Valoración de las provisiones de siniestros" 2010/2011
- 115.- Rosa Viciano García: "Banca-Seguros. Evolución, regulación y nuevos retos" 2007/2008
- 116.- Ramón Vidal Escobosa: "El baremo de Daños Personales en el Seguro de Automóviles" 2009/2010
- 117.- Tomás Wong-Kit Ching: "Análisis del Reaseguro como mitigador del capital de riesgo" 2008/2009
- 118.- Yibo Xiong: "Estudio del mercado chino de Seguros: La actualidad y la tendencia" 2005/2006
- 119.- Beatriz Bernal Callizo: "Póliza de Servicios Asistenciales" 2003/2004
- 120.- Marta Bové Badell: "Estudio comparativo de evaluación del Riesgo de Incendio en la Industria Química" 2003/2004
- 121.- Ernest Castellón Teixidó: "La edificación. Fases del proceso, riesgos y seguros" 2004/2005
- 122.- Sandra Clusella Giménez: "Gestió d'Actius i Passius. Inmunització Financera" 2004/2005
- 123.- Miquel Crespí Argemí: "El Seguro de Todo Riesgo Construcción" 2005/2006
- 124.- Yolanda Dengra Martínez: "Modelos para la oferta de seguros de Hogar en una Caja de Ahorros" 2007/2008
- 125.- Marta Fernández Ayala: "El futuro del Seguro. Bancaseguros" 2003/2004

- 126.- Antonio Galí Isus: "Inclusión de las Energías Renovables en el sistema Eléctrico Español" 2009/2010
- 127.- Gloria Gorbea Bretones: "El control interno en una entidad aseguradora" 2006/2007
- 128.- Marta Jiménez Rubio: "El procedimiento de tramitación de siniestros de daños materiales de automóvil: análisis, ventajas y desventajas" 2008/2009
- 129.- Lorena Alejandra Libson: "Protección de las víctimas de los accidentes de circulación. Comparación entre el sistema español y el argentino" 2003/2004
- 130.- Mario Manzano Gómez: "La responsabilidad civil por productos defectuosos. Solución aseguradora" 2005/2006
- 131.- Àlvar Martín Botí: "El Ahorro Previsión en España y Europa. Retos y Oportunidades de Futuro" 2006/2007
- 132.- Sergio Martínez Olivé: "Construcción de un modelo de previsión de resultados en una Entidad Aseguradora de Seguros No Vida" 2003/2004
- 133.- Pilar Miracle Vázquez: "Alternativas de implementación de un Departamento de Gestión Global del Riesgo. Aplicado a empresas industriales de mediana dimensión" 2003/2004
- 134.- María José Morales Muñoz: "La Gestión de los Servicios de Asistencia en los Multirriesgo de Hogar" 2007/2008
- 135.- Juan Luis Moreno Pedroso: "El Seguro de Caución. Situación actual y perspectivas" 2003/2004
- 136.- Rosario Isabel Pastrana Gutiérrez: "Creació d'una empresa de serveis socials d'atenció a la dependència de les persones grans enfocada a productes d'assegurances" 2007/2008
- 137.- Joan Prat Rifà: "La Previsió Social Complementaria a l'Empresa" 2003/2004
- 138.- Alberto Sanz Moreno: "Beneficios del Seguro de Protección de Pagos" 2004/2005
- 139.- Judith Safont González: "Efectes de la contaminació i del estils de vida sobre les assegurances de salut i vida" 2009/2010
- 140.- Carles Soldevila Mejías: "Models de gestió en companyies d'assegurances. Outsourcing / In-sourcing" 2005/2006
- 141.- Olga Torrente Pascual: "IFRS-19 Retribuciones post-empleo" 2003/2004
- 142.- Annabel Roig Navarro: "La importancia de las mutualidades de previsión social como complementarias al sistema público" 2009/2010
- 143.- José Angel Ansón Tortosa: "Gerencia de Riesgos en la Empresa española" 2011/2012
- 144.- María Mercedes Bernués Burillo: "El permiso por puntos y su solución aseguradora" 2011/2012
- 145.- Sònia Beulas Boix: "Prevención del blanqueo de capitales en el seguro de vida" 2011/2012
- 146.- Ana Borràs Pons: "Teletrabajo y Recursos Humanos en el sector Asegurador" 2011/2012
- 147.- María Asunción Cabezas Bono: "La gestión del cliente en el sector de bancaseguros" 2011/2012
- 148.- María Carrasco Mora: "Matching Premium. New approach to calculate technical provisions Life insurance companies" 2011/2012
- 149.- Eduard Huguet Palouzie: "Las redes sociales en el Sector Asegurador. Plan social-media. El Community Manager" 2011/2012
- 150.- Laura Monedero Ramírez: "Tratamiento del Riesgo Operacional en los 3 pilares de Solvencia II" 2011/2012
- 151.- Salvador Obregón Gomá: "La Gestión de Intangibles en la Empresa de Seguros" 2011/2012
- 152.- Elisabet Ordóñez Somolinos: "El sistema de control Interno de la Información Financiera en las Entidades Cotizadas" 2011/2012
- 153.- Gemma Ortega Vidal: "La Mediación. Técnica de resolución de conflictos aplicada al Sector Asegurador" 2011/2012
- 154.- Miguel Ángel Pino García: "Seguro de Crédito: Implantación en una aseguradora multirramo" 2011/2012
- 155.- Genevieve Thibault: "The Costumer Experience as a Sorce of Competitive Advantage" 2011/2012

- 156.- Francesc Vidal Bueno: "La Mediación como método alternativo de gestión de conflictos y su aplicación en el ámbito asegurador" 2011/2012
- 157.- Mireia Arenas López: "El Fraude en los Seguros de Asistencia. Asistencia en Carretera, Viaje y Multirriesgo" 2012/2013
- 158.- Lluís Fernández Rabat: "El proyecto de contratos de Seguro-IFRS4. Expectativas y realidades" 2012/2013
- 159.- Josep Ferrer Arilla: "El seguro de decesos. Presente y tendencias de futuro" 2012/2013
- 160.- Alicia García Rodríguez: "El Cuadro de Mando Integral en el Ramo de Defensa Jurídica" 2012/2013
- 161.- David Jarque Solsona: "Nuevos sistemas de suscripción en el negocio de vida. Aplicación en el canal bancaseguros" 2012/2013
- 162.- Kamal Mustafá Gondolbeu: "Estrategias de Expansión en el Sector Asegurador. Matriz de Madurez del Mercado de Seguros Mundial" 2012/2013
- 163.- Jordi Núñez García: "Redes Periciales. Eficacia de la Red y Calidad en el Servicio" 2012/2013
- 164.- Paula Núñez García: "Benchmarking de Autoevaluación del Control en un Centro de Siniestros Diversos" 2012/2013
- 165.- Cristina Riera Asensio: "Agregadores. Nuevo modelo de negocio en el Sector Asegurador" 2012/2013
- 166.- Joan Carles Simón Robles: "Responsabilidad Social Empresarial. Propuesta para el canal de agentes y agencias de una compañía de seguros generalista" 2012/2013
- 167.- Marc Vilardebó Miró: "La política de inversión de las compañías aseguradoras ¿Influirá Solvencia II en la toma de decisiones?" 2012/2013
- 168.- Josep María Bertrán Aranés: "Segmentación de la oferta aseguradora para el sector agrícola en la provincia de Lleida" 2013/2014
- 169.- María Buendía Pérez: "Estrategia: Formulación, implementación, valoración y control" 2013/2014
- 170.- Gabriella Fernández Andrade: "Oportunidades de mejora en el mercado de seguros de Panamá" 2013/2014
- 171.- Alejandro Galcerán Rosal: "El Plan Estratégico de la Mediación: cómo una Entidad Aseguradora puede ayudar a un Mediador a implementar el PEM" 2013/2014
- 172.- Raquel Gómez Fernández: "La Previsión Social Complementaria: una apuesta de futuro" 2013/2014
- 173.- Xoan Jovaní Guiral: "Combinaciones de negocios en entidades aseguradoras: una aproximación práctica" 2013/2014
- 174.- Àlex Lansac Font: "Visión 360 de cliente: desarrollo, gestión y fidelización" 2013/2014
- 175.- Albert Llambrich Moreno: "Distribución: Evolución y retos de futuro: la evolución tecnológica" 2013/2014
- 176.- Montserrat Pastor Ventura: "Gestión de la Red de Mediadores en una Entidad Aseguradora. Presente y futuro de los agentes exclusivos" 2013/2014
- 177.- Javier Portalés Pau: "El impacto de Solvencia II en el área de TI" 2013/2014
- 178.- Jesús Rey Pulido: "El Seguro de Impago de Alquileres: Nuevas Tendencias" 2013/2014
- 179.- Anna Solé Serra: "Del cliente satisfecho al cliente entusiasmado. La experiencia cliente en los seguros de vida" 2013/2014
- 180.- Eva Tejedor Escorihuela: "Implantación de un Programa Internacional de Seguro por una compañía española sin sucursales o filiales propias en el extranjero. Caso práctico: Seguro de Daños Materiales y RC" 2013/2014
- 181.- Vanesa Cid Pijuan: "Los seguros de empresa. La diferenciación de la mediación tradicional" 2014/2015.
- 182.- Daniel Ciprés Tiscar: "¿Por qué no arranca el Seguro de Dependencia en España?" 2014/2015.
- 183.- Pedro Antonio Escalona Cano: "La estafa de Seguro. Creación de un Departamento de Fraude en una entidad aseguradora" 2014/2015.

- 184.- Eduard Escardó Lleixà: "Análisis actual y enfoque estratégico comercial de la Bancaseguros respecto a la Mediación tradicional" 2014/2015.
- 185.- Marc Esteve Grau: "Introducción del Ciber Riesgo en el Mundo Asegurador" 2014/2015.
- 186.- Paula Fernández Díaz: "La Innovación en las Entidades Aseguradoras" 2014/2015.
- 187.- Alex Lleyda Capell: "Proceso de transformación de una compañía aseguradora enfocada a producto, para orientarse al cliente" 2014/2015.
- 188.- Oriol Petit Salas: "Creación de Correduría de Seguros y Reaseguros S.L. Gestión Integral de Seguros" 2014/2015.
- 189.- David Ramos Pastor: "Big Data en sectores Asegurador y Financiero" 2014/2015.
- 190.- Marta Raso Cardona: "Comoditización de los seguros de Autos y Hogar. Diferenciación, fidelización y ahorro a través de la prestación de servicios" 2014/2015.
- 191.- David Ruiz Carrillo: "Información de clientes como elemento estratégico de un modelo asegurador. Estrategias de Marketing Relacional/CRM/Big Data aplicadas al desarrollo de un modelo de Bancaseguros" 2014/2015.
- 192.- María Torrent Caldas: "Ahorro y planificación financiera en relación al segmento de jóvenes" 2014/2015.
- 193.- Cristian Torres Ruiz: "El seguro de renta vitalicia. Ventajas e inconvenientes" 2014/2015.
- 194.- Juan José Trani Moreno: "La comunicación interna. Una herramienta al servicio de las organizaciones" 2014/2015.
- 195.- Alberto Yebra Yebra: "El seguro, producto refugio de las entidades de crédito en épocas de crisis" 2014/2015.
- 196.- Jesús García Riera: "Aplicación de la Psicología a la Empresa Aseguradora" 2015/2016
- 197.- Pilar Martínez Beguería: "La Función de Auditoría Interna en Solvencia II" 2015/2016
- 198.- Ingrid Nicolás Fargas: "El Contrato de Seguro y su evolución hasta la Ley 20/2015 LOSSEAR. Hacia una regulación más proteccionista del asegurado" 2015/2016
- 199.- María José Páez Reigosa: "Hacia un nuevo modelo de gestión de siniestros en el ramo de Defensa Jurídica" 2015/2016
- 200.- Sara Melissa Pinilla Vega: "Auditoría de Marca para el Grupo Integra Seguros Limitada" 2015/2016
- 201.- Teresa Repollés Llecha: "Optimización del ahorro a través de soluciones integrales. ¿cómo puede la empresa ayudar a sus empleados? 2015/2016
- 202.- Daniel Rubio de la Torre: "Telematics y el seguro del automóvil. Una nueva póliza basada en los servicios" 2015/2016
- 203.- Marc Tarragó Diego: "Transformación Digital. Evolución de los modelos de negocio en las compañías tradicionales" 2015/2016
- 204.- Marc Torrents Fábregas: "Hacia un modelo asegurador peer-to-peer. ¿El modelo asegurador del futuro?" 2015/2016
- 205.- Inmaculada Vallverdú Coll: "Fórmulas modernas del Seguro de Crédito para el apoyo a la empresa: el caso español" 2015/2016
- 206.- Cristina Alberch Barrio: "Seguro de Crédito. Gestión y principales indicadores" 2016/2017
- 207.- Ian Bachs Millet: "Estrategias de expansión geográfica de una entidad aseguradora para un mercado específico" 2016/2017
- 208.- Marta Campos Comas: "Externalización del servicio de asistencia" 2016/2017
- 209.- Jordi Casas Pons: "Compromisos por pensiones. Hacia un nuevo modelo de negociación colectiva" 2016/2017
- 210.- Ignacio Domenech Guillén: "El seguro del automóvil para vehículos sostenibles, autónomos y conectados" 2016/2017
- 211.- María Luisa Fernández Gómez: "Adquisiciones de Carteras de Seguros y Planes de Pensiones" 2016/2017
- 212.- Diana Heman Hasbach: "¿Podrán los Millennials cobrar pensión?: una aplicación al caso de México" 2016/2017

- 213.- Sergio López Serrano: "El impacto de los Ciberriesgos en la Gerencia de Riesgos Tradicional" 2016/2017
- 214.- Jordi Martí Bernaus: "Dolencias preexistentes en el seguro de Salud: exclusiones o sobreprimas" 2016/2017
- 215.- Jéssica Martínez Ordóñez: "Derecho al honor de las personas jurídicas y reputación online" 2016/2017
- 216.- Raúl Monjo Zapata: "La Función de Cumplimiento en las Entidades Aseguradoras" 2016/2017
- 217.- Francisco José Muñoz Guerrero: "Adaptación de los Productos de Previsión al Ciclo de Vida" 2016/2017
- 218.- Mireia Orenes Esteban: "Crear valor mediante la gestión de siniestros de vida" 2016/2017
- 219.- Oscar Pallisa Gabriel: "Big Data y el sector asegurador" 2016/2017
- 220.- Marc Parada Ricart: "Gerencia de Riesgos en el Sector del Transporte de Mercancías" 2016/2017
- 221.- Xavier Pérez Prado: "Análisis de la mediación en tiempos de cambio. Debilidades y fortalezas. Una visión de futuro" 2016/2017
- 222.- Carles Pons Garulo: "Solvencia II: Riesgo Catastrófico. Riesgo Antropógeno y Reaseguro en el Seguro de Daños Materiales" 2016/2017
- 223.- Javier Pulpillo López: "El Cuadro de Mando Integral como herramienta de gestión estratégica y retributiva" 2016/2017