



UNIVERSIDAD DE DEUSTO

DESARROLLO DE UN MODELO DE COMPORTAMIENTO DEL VISITANTE VIRTUAL EN EL MARCO DE LOS PROVEEDORES DE PRIMER NIVEL EN EL SECTOR DE LA DISTRIBUCIÓN

Tesis doctoral presentada por Iñigo Castánder Martín
dentro del Programa de Doctorado en Ingeniería para la Sociedad de la
Información y Desarrollo Sostenible
Dirigida por Dra. Ainhoa Serna Nocedal



UNIVERSIDAD DE DEUSTO

DESARROLLO DE UN MODELO DE COMPORTAMIENTO DEL VISITANTE VIRTUAL EN EL MARCO DE LOS PROVEEDORES DE PRIMER NIVEL EN EL SECTOR DE LA DISTRIBUCIÓN

Tesis doctoral presentada por Iñigo Castánder Martín
dentro del Programa de Doctorado en Ingeniería para la Sociedad de la
Información y Desarrollo Sostenible
Dirigida por Dra. Ainhoa Serna Nocedal

El doctorando

La directora

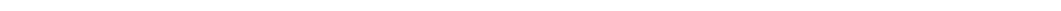
La directora

Bilbao, mayo 2019

Desarrollo de un modelo de comportamiento del visitante virtual en el marco de los proveedores de primer nivel en el sector de la distribución

Autor: Iñigo Castánder Martín

Directora: Ainhoa Serna Nocedal



*Dedicado especialmente a mis padres (Iñaki y Rosi),
a mi hermana Bea
y a mi novia Eva.*

Siempre dos hay. No más, no menos. Un maestro, y un aprendiz.

— *Maestro Yoda (Star Wars, Episodio I)*

Resumen

Esta tesis doctoral tiene como objetivo principal desarrollar un modelo de comportamiento del visitante virtual en el marco de los proveedores de primer nivel (los más cercanos en la cadena de suministro a los distribuidores) en el sector de la distribución, ya que no se ha detectado ningún modelo en la revisión de la literatura sobre el comportamiento de los visitantes virtuales en tales proveedores. Por ello, el fin último será, por un lado, aportar a la comunidad científica el modelo que cubra el *gap* identificado. Por otro lado, la transferencia para explotación empresarial del *software* y conocimiento descubierto a través de esta investigación.

La metodología que se plantea para llevar el proyecto a cabo tiene un enfoque cualitativo, siendo el tipo de estudio exploratorio con casos de estudio. Además, los datos del proyecto se obtendrán mediante diferentes técnicas de recolección de datos: 1) *Web Analytics* y 2) observaciones tras los contactos con el caso de estudio.

La contribución y/o novedad que aporta este proyecto de investigación es la generación de conocimiento sobre el comportamiento de los visitantes virtuales a través de: a) la identificación de una cadena de suministro física y virtual, estableciendo el papel de los diversos agentes que interceden en la misma y b) la definición de las ramas de análisis de comportamiento del visitante virtual y de sus factores, para estudiar la situación actual y las predicciones obtenidas del análisis de *Web Mining*. Posibilitando así, la generación de la estrategia analítica en base a determinadas pautas, para que el proveedor de primer nivel pueda emplear lo aprendido sobre el comportamiento del visitante virtual, en la evolución del sitio web y otras implementaciones tecnológicas. De

esta forma, el modelo resultante podrá ser generalizado analíticamente a otros sectores y casos de estudio.

Palabras clave: visitantes virtuales, sector distribución, proveedores de primer nivel, modelo de comportamiento, analítica web

Abstract

The main objective of this doctoral thesis is to develop a virtual visitor behavior model within the framework of the tier 1 suppliers (the closest ones in the supply chain to distributors) in the distribution sector, since it has not been detected no model in the literature review on the behavior of virtual visitors in such providers. Therefore, the ultimate goal will be, on the one hand, to provide the scientific community with the model that covers the identified gap. On the other hand, the transfer for business exploitation of software and knowledge discovered through this research.

The methodology proposed to carry out the project has a qualitative approach, being the type of exploratory study with case studies. In addition, project data will be obtained through different data collection techniques: 1) Web Analytics and 2) observations after contacts with the case study.

The contribution and/or novelty that this research project contributes is the generation of knowledge about the behavior of virtual visitors through: a) the identification of a physical and virtual supply chain, establishing the role of the various agents that intercede in the same and b) the definition of the branches of behavior analysis of the virtual visitor and its factors, to study the current situation and the predictions obtained from the Web Mining analysis. Making possible, the generation of the analytical strategy based on certain guidelines, so that the first level provider can use what they have learned about the behavior of the virtual visitor, in the evolution of the website and other technological implementations. In this way, the resulting model can be generalized analytically to other sectors and case studies.

Keywords: virtual visitors, distribution sector, tier 1 suppliers, behavior model, web analytics

Laburpena

Doktorego tesi honen helburu nagusia bisitari birtualen portaera eredu bat garatzea da, lehen mailako hornitzairen esparruan (hornikuntza katean banatzaileetik hurbilen daudenak), banaketa sektorean. Ez baita detektatu literaturaren berrikuspenean bisitari birtualen portaerari buruzko eredu bat, hornitzaire horietan. Hori dela eta, azken helburua izango da, alde batetik, komunitate zientifikora identifikatutako hutsunea betetzen duen eredua ematea. Bestalde, ikerketa honen bitartez garatutako software eta ezagutza enpresen ustiapenerako transferentzia egitea.

Proiektua burutzeko proposatutako metodologia ikuspegi kualitatiboa da eta azterketa mota exploratorioa izango da, kasu azterketekin. Horrez gain, proiektuaren datuak datu bilketako hainbat tekniken bidez lortuko dira: 1) *Web Analytics* bitartez eta 2) kasu azterketarekin harremanetan jarri ondoren egindako oharrekin.

Ikerketa-proiektu honek ematen duen ekarpena eta/edo berritasuna, bisitari birtualen portaerari buruzko ezagutza sortzea da: a) hornikuntzakate fisiko eta birtualaren identifikazioaren bidez, bertan jarduten duten eragile guztien papera finkatuz. b) bisitari birtualaren portaeraren analisi atalak eta beren faktoreen definizioaren bidez, uneko egoera eta *Web Mining* azterketaren bitartez lortutako iragarpenak aztertze-ko. Zenbait jarraibideetan oinarritutako estrategia analitikoa sortzea ahalbidetuz, lehen mailako hornitzaireak bisitari birtualaren portaerari buruz ikasitakoa erabil dezan, webgunearen bilakaeran eta beste implementazio teknologikoetan. Horrela, sortzen den eredua beste sektore eta kasuen azterketarako analitikoki orokortu daiteke.

Hitz gakoak: bisitari birtualak, banaketa sektorea, lehen mailako hor-nitzaireak, portaera eredua, web analitika

Agradecimientos

Sirvan las siguientes líneas como agradecimiento a todas las personas que me han apoyado durante mi carrera profesional y concretamente en este proyecto de investigación.

En primer lugar, quiero dar las gracias a i68 y a Isabel Busto en particular, por darme la oportunidad de poder llevar a cabo este proyecto de investigación. A pesar de todo, para mí es muy importante recalcar que Isabel ha sido la tutora de empresa de este proyecto. Nunca olvidaré las numerosas reuniones y sobre todo los bocetos y borradores en papel escritos.

En segundo lugar, agradecerle a Ainhoa Serna su insistencia y apoyo durante todos estos años de proyecto. En el cuál, sin la participación de ORBEA S.Coop., hubiese sido difícil encontrar un caso de estudio tan amplio, por lo tanto, gracias de corazón a Alberto y a Daniel.

En tercer lugar, quiero dar las gracias a todos los compañeros de i68 y concretamente de la Unidad, que me han ayudado con el proyecto, e incluso a aquellos que me preguntaban sobre el mismo, aun no entendiendo absolutamente nada de lo que les explicaba.

En cuarto lugar, a mis amigos y familiares, que llevan preguntándome año tras año cuando voy a terminar la tesis. No obstante, siempre he percibido por parte de ellos una sensación de orgullo y de mérito, por el mero hecho de que alguien cercano a ellos, haya llegado hasta este punto. ¡Gracias!

No me quiero olvidar también, de mi novia, la cual ha sufrido todos los buenos y malos momentos que he pasado durante este proyecto. Gracias, por tu comprensión, tus consejos y por tu ayuda a este chico de ciencias que sabe algo de letras.

Por último, gracias a mi hermana y a mis padres. Solamente con mirarles sé lo que se alegran de que haya llegado hasta aquí. A pesar de los logros académicos que he conseguido, lo que nunca voy a olvidar es todo lo que mis padres me han enseñado y los valores de: esfuerzo, sacrificio, trabajo y de no rendirse nunca, que me han transmitido. Pero sobre todo, por su apoyo y su fuerza cuando no hacía más que fallar y solo ellos estaban para ayudarme a superar dichos momentos. Gracias de verdad, me siento un privilegiado por teneros a mi lado.

Eskerrik asko,

Iñigo

mayo 2019

Índice general

Índice de figuras	xv
Índice de tablas	xvii
1 Introducción	1
1.1 Contexto	1
1.2 Propósito y objetivos del proyecto	3
1.3 Planteamiento del problema	5
1.4 Contribución	7
1.5 Preguntas de Investigación	8
1.6 Metodología	10
1.7 Estructura de la investigación	10
1.8 Proyectos, publicaciones y transferencia a empresa	11
1.8.1 Proyectos	11
1.8.2 Publicaciones	13
2 Estado del Arte	15
2.1 El sector de la distribución	15
2.1.1 Estructura	16
2.1.2 Características y necesidades del sector	18
2.1.3 Comparativa entre las cadenas de suministro física y virtual	25
2.1.4 Recorrido de los componentes de la cadena: desde la pers- pectiva física a la virtual	26

ÍNDICE GENERAL

2.1.5	Retrospectiva: de la integración al comportamiento en los canales de venta	31
2.2	Cadena de suministro virtual	38
2.2.1	Escenarios de relación	39
2.2.2	Consideraciones sobre las ramas de análisis de comportamiento	40
2.3	Conocimiento del comportamiento del visitante virtual a través del <i>Web Analytics</i>	41
2.3.1	Obtención y acceso a las fuentes de datos del comportamiento	42
2.3.2	Integración de los datos de comportamiento	48
2.3.3	Ánalisis de los datos de comportamiento	51
2.3.3.1	Método Estadístico	51
2.3.3.2	Data Mining	52
2.3.3.3	Web Mining	53
2.3.4	Técnicas de <i>Data Mining</i> disponibles en el <i>Web Mining</i> . .	55
2.3.4.1	ANN (<i>Artificial Neural Network</i>) & SOM (<i>Self-Organizing Mapping</i>)	55
2.3.4.2	Clustering	56
2.3.4.3	LDA (<i>Linear Discriminant Analysis</i>) & LR (<i>Logistic Regression</i>)	56
2.3.4.4	CART (<i>Classification And Regression Tree</i>) & RF (<i>Random Forest</i>)	57
2.3.4.5	AR (<i>Association Rules</i>)	57
2.3.4.6	SVM (<i>Support Vector Machines</i>)	58
2.3.4.7	GA (<i>Genetic Algorithms</i>)	58
3	Metodología de la investigación	61
3.1	Método de investigación	62
3.1.1	Enfoques de investigación	62
3.1.2	Propuesta de investigación	63
3.2	Diseño de la investigación	67
3.2.1	Preguntas de investigación	67
3.2.1.1	Preguntas asociadas al primer objetivo específico	68

ÍNDICE GENERAL

3.2.1.2	Preguntas asociadas al segundo objetivo específico	68
3.2.1.3	Preguntas asociadas al tercer objetivo específico	69
3.2.1.4	Preguntas asociadas al cuarto objetivo específico	69
3.2.2	Definición del contexto	70
3.2.3	Criterios de selección del caso de estudio	71
3.2.4	Metodología de recolección de datos	71
3.2.4.1	Origen de datos	72
3.2.4.2	Selección de las herramientas de <i>Web Tracking</i> .	73
3.2.5	Herramienta de análisis de datos	74
3.2.6	Trabajo de campo	74
3.3	Fiabilidad y validez de esta investigación	76
4	Análisis de la evidencia empírica	79
5	Generación del modelo	81
6	Conclusiones, limitaciones y líneas futuras	83
6.1	Revisión de los objetivos	83
6.2	Revisión de las preguntas de investigación	89
6.3	Conclusiones	102
6.3.1	Conclusión global	102
6.3.2	Conclusión sobre la metodología aplicada	103
6.3.3	Conclusiones sobre los resultados	103
6.4	Aportaciones	104
6.5	Limitaciones y líneas futuras	105
6.5.1	Limitaciones	105
6.5.2	Líneas futuras	105
A	Visualizaciones del Análisis Web Mining	107
Bibliografía		108
		108
Glosario		121

Índice de figuras

1.1	Objetivos general y específicos de la Investigación	5
2.1	Modelo Push-Pull en la cadena de valor	17
2.2	Porcentaje de cuota de mercado de las cinco primeras empresas según el país europeo	22
2.3	Evolución del peso de los segmentos de consumidor (sobre la población total)	24
2.4	Comparativa entre las cadenas de suministro física y virtual	26
2.5	Escenarios de Relación	39
2.6	Cadena de Valor del sector de la distribución	40
2.7	Estudio de variables - Recorrido desde las genéricas hasta las de comercio electrónico de <i>Google Analytics</i> y <i>Piwik</i>	47
2.8	Estudio de variables - Relaciones entre métricas web y de <i>e-business</i>	48
2.9	Cuadrante Mágico de <i>Gartner</i> para BI & Analytics Platforms 2018	50
2.10	Esquema del <i>Web Mining</i>	54
3.1	Esquema de resumen de comparativa de enfoques de investigación	63
3.2	Esquema general de la metodología	64
3.3	Ánálisis de diversas estrategias para una investigación cualitativa .	65
3.4	Esquema de resumen de la propuesta de metodología	67
3.5	Cronograma de trabajo de campo	75
6.1	Propuesta de estructura global de la cadena de suministro física y virtual (Recapitulación)	85
6.2	Flujo de la visita de los visitantes virtuales (Recapitulación)	86

ÍNDICE DE FIGURAS

6.3	Esquema de la estrategia analítica (Recapitulación)	89
6.4	Cadena de suministro física (Recapitulación)	91
6.5	Cadena de suministro física + virtual (Recapitulación)	92
6.6	Ramas de análisis de comportamiento del visitante virtual por es- cenario de relación. (Recapitulación)	94

Índice de tablas

1.1	Preguntas de Investigación por Objetivos Específicos	9
2.1	Características del sector de la distribución estatal	18
2.2	Recorrido desde la parte física hasta la virtual en la cadena de suministro del sector de la distribución.	31
2.3	Retrospectiva de la integración al comportamiento en los canales de venta en la cadena de suministro del sector de la distribución. .	38
2.4	Componentes de comercio electrónico para los sistemas de recogida de información	43
2.5	Comparativa de las herramientas de <i>Web Tracking</i> a utilizar para el origen de <i>Web Analytics</i>	45
3.1	Proceso del Origen de Datos	72
3.2	Origen de datos que da respuesta a cada pregunta de investigación	73

Vas hacia el camino y si no cuidas tus pasos, no sabes hacia dónde te arrastrarán.

Bilbo Bolsón - El señor de los anillos (La comunidad del anillo)

CAPÍTULO

1

Introducción

En este capítulo se contextualiza la investigación realizada sobre el comportamiento del visitante virtual en el sector de la distribución, en el marco de los proveedores de primer nivel. Después, se expone el propósito y objetivos específicos de esta tesis, así como el planteamiento del problema, el caso de estudio llevado a cabo y la contribución tanto en el ámbito científico-académico, como empresarial. En este último, es importante recalcar la relación directa de esta tesis con el proyecto industrial desarrollado en i68, por el cual se ha solicitado la mención industrial. Además, se introduce brevemente a la metodología aplicada a través de la generación de preguntas de investigación, que más tarde en el capítulo 3 se detallarán en profundidad. Asimismo, se describen los capítulos que la conforman. Por último, para finalizar se detallan los proyectos y publicaciones desarrollados durante la elaboración de la tesis.

1.1 Contexto

La oportunidad que fue la irrupción de la era digital en la evolución de los sistemas de información (Belk, Papatheocharous, Germanakos, & Samaras, 2013; Indra, 2015) y el impacto de Internet (Best Education, 2013) permitió la evolución de los canales de información y del comportamiento de las personas respecto a estos

Capítulo 1. Introducción

sistemas, y sobre todo a lo que supuso su relación con Internet (Rubio Gil, 2010; Sánchez Valle & De Frutos Torres, 2012; Fernández Morales, 2014; López Maleda, 2014; Cloquell Lozano, 2015). Además, a nivel empresarial esta evolución digital favoreció a los sistemas de información de gestión y desde una perspectiva extendida, a aquellas herramientas que permitieron interactuar con personal externo a la organización. De hecho, esta situación generó diversas relaciones entre agentes que, en el caso de las orientadas al comercio electrónico y la de la gestión de las relaciones con los clientes, generaron la posibilidad de ahondar en el conocimiento del comportamiento en el visitante virtual.

Asimismo, el crecimiento del comercio electrónico (Labrecque, Mathwick, Novak, Hofacker, et al., 2013; ONTSI, 2013, 2015), en base a determinadas plataformas (INE, 2013) que interactúan o son parte de los sistemas de información de gestión empresarial comentados anteriormente, fue el motivo principal por el cual se escogió al sector de distribución de España. Concretamente, se puso el foco en los proveedores de primer nivel, ya que era en éstos donde se detectó un *gap* respecto a los otros agentes de la cadena de valor del sector, en lo relativo a las investigaciones realizadas respecto, al comportamiento de los visitantes virtuales (Castánder, Serna, & Busto, 2017).

Por ello y antes de entrar en el propósito del proyecto, a continuación se van dar unos datos concretos sobre la investigación llevada a cabo. En primer lugar, exponer que el caso de estudio se realiza con la empresa ORBEA S.Coop. tras un análisis minucioso de las empresas del sector de la distribución en el marco de los proveedores de primer nivel.

Los datos para el caso de estudio se han recogido durante 21 meses. Exactamente, ha abarcado un periodo de datos que transcurre desde enero de 2017 hasta septiembre de 2018, con un origen de datos formado por dos fuentes de datos diferentes:

1. un conjunto de datos sobre las visitas que han realizado los visitantes virtuales al sitio web del caso de estudio, a través de las técnicas de *Web Analytics*.
2. las observaciones realizadas tras las conversaciones con el caso de estudio.

1.2 Propósito y objetivos del proyecto

Los métodos de análisis de datos utilizados han sido el análisis heurístico y la aplicación de técnicas de *Web Mining* mediante una herramienta de visualización y análisis de datos. De hecho, se ha manejado un volumen de datos de alrededor de 9 millones de registros. Finalmente, la metodología que se ha utilizado para este proyecto ha sido la cualitativa exploratoria con estudio de un único caso.

1.2 Propósito y objetivos del proyecto

El **propósito de esta investigación** ha sido generar un modelo sobre el comportamiento del visitante virtual en el sector de la distribución, a través del conocimiento obtenido en el análisis de la evidencia empírica. El modelo que se ha generado dispone de un conjunto de pautas, englobadas en la llamada estrategia analítica, que pueden ser de utilidad para las organizaciones que son proveedoras de primer nivel en el sector de la distribución.

Asimismo, es fundamental remarcar que este proyecto tiene como uno de los objetivos principales crear una nueva línea de producto que redundará en una diversificación de la oferta actual de i68. Incluyendo la mejora del ERP IZARO (producto de i68), ya que este proyecto va a ayudar a un crecimiento, desarrollando un nuevo producto que complementa y refuerza el impacto del nivel extendido del propio ERP. Además, también va a impactar en otros proyectos importantes de la unidad de I+D+i, al poder aplicar el conocimiento descubierto sobre los visitantes virtuales en el comercio electrónico. De hecho, cabe recalcar que se ha realizado una transferencia del modelo y del software implementado al sector empresarial y concretamente a Informática 68 (Grupo Zucchetti) una empresa referente en su sector. Permitiendo así una aplicación práctica, efectiva y real de los conocimientos profesionales de la investigación.

El interés por esta investigación se fundamentó en la inexistencia de un modelo de conocimiento sobre el comportamiento del visitante virtual en el marco de los proveedores de primer nivel en el sector de la distribución. En consecuencia, se implementó dicho modelo de conocimiento en base a las preguntas de investigación realizadas. Este modelo, a partir del amplio y profundo

Capítulo 1. Introducción

análisis (con diferentes escenarios) de un caso de estudio, definiría un conjunto de pautas de actuación englobadas en una estrategia analítica para dichos proveedores.

El fin último de este modelo, ha sido facilitar a los agentes de dicho caso de estudio, la comprensión sobre los visitantes virtuales, a través de la definición de un conjunto de pautas que les ayudasen a generar una estrategia analítica, tanto para valorar desarrollos tecnológicos futuros, como para el desarrollo de estrategias comerciales y de marketing futuras.

Este modelo podría ser de interés, para diferentes perfiles de puesto dentro de la empresa objeto del caso de estudio, como por ejemplo: los responsables de las plataformas de comercio electrónico, responsables de la gestión de las relaciones con los clientes, etc.

Para alcanzar este propósito u objetivo general del proyecto se definieron un conjunto de objetivos específicos, aplicados al marco de los proveedores de primer nivel dentro del sector de la distribución.

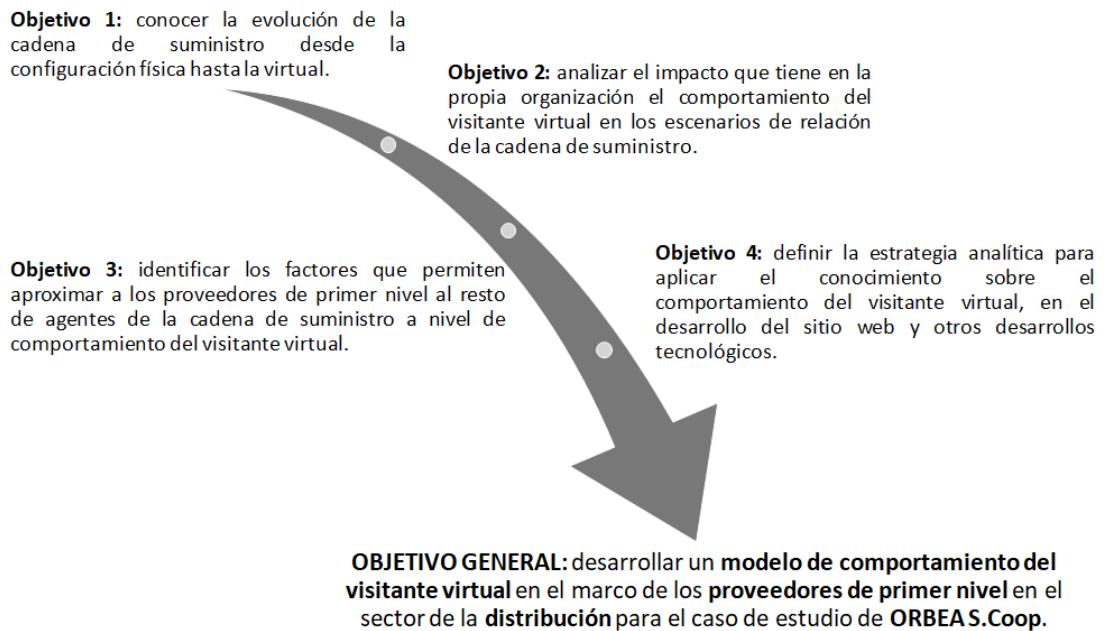
Los objetivos específicos son los siguientes:

- **Objetivo 1:** conocer la evolución de la cadena de suministro desde la configuración física hasta la virtual.
- **Objetivo 2:** analizar el impacto que tiene en la propia organización, el comportamiento del visitante virtual en los escenarios de relación de la cadena de suministro.
- **Objetivo 3:** identificar los factores que permiten aproximar a los proveedores de primer nivel, al resto de agentes de la cadena de suministro a nivel de comportamiento del visitante virtual.
- **Objetivo 4:** definir la estrategia analítica para aplicar el conocimiento sobre el comportamiento del visitante virtual, en el desarrollo del sitio web y otros desarrollos tecnológicos.

1.3 Planteamiento del problema

A continuación, se muestra un esquema con los objetivos específicos descritos anteriormente:

Figura 1.1: Objetivos general y específicos de la Investigación



Fuente: Elaboración propia

1.3 Planteamiento del problema

La coyuntura económica y social que se venía arrastrando en los últimos años, junto con el impacto que suponía la utilización de las Tecnologías de la Información y Comunicación (TIC) (Belk et al., 2013) e Internet (Best Education, 2013; Zhang, Zhao, Cheung, & Lee, 2014) tuvo como consecuencia que muchos sectores empresariales de la sociedad (industria (fabricación, automoción,...), servicios, etc.) se vieran inmersos o en el comienzo de procesos de reconversión.

Uno de ellos fue el sector del gran consumo y la distribución, el cual se pudo ver fortalecido y a su vez pudo haber agilizado la recuperación económica, si los consumidores hubieran tenido mayor poder adquisitivo (Bonmatí, 2015). De esta manera, se pudo favorecer el aumento del número de personas que utilizaban el

Capítulo 1. Introducción

comercio electrónico que tuvo cada vez un mayor crecimiento (Labrecque et al., 2013; ONTSI, 2013, 2015) y peso en el propio sector. No hay que olvidar que todo ello propició un cambio en el estilo de vida de las personas y esto a su vez, junto con lo comentado anteriormente, influyó sobre la propia distribución. Todo un reto que puso el foco en aumentar la presencia en un mercado con mayor competitividad, menos superficies y menos clientes (Indra, 2015). Un sector que como se describe más adelante tiene sus características y necesidades (ver apartado 2.1.2) que junto con la comunidad de personas que están formando los visitantes virtuales, debido a la utilización de herramientas de comercio electrónico, dan pie a la oportunidad de ahondar en el conocimiento sobre estos visitantes. Es decir, conocer como se comportan a través de estas plataformas y por consiguiente cómo puede afectar al sector de la distribución.

Para **profundizar en el conocimiento de sobre los visitantes virtuales**, este proyecto de investigación, **se ha llevado a cabo sobre los diferentes escenarios de relación, con los visitantes/clientes virtuales que disponía la empresa ORBEA S.Coop.**, la cual, ofrecía una amplia casuística en la cadena de valor (suministro), ya que disponía de sitio web corporativo, del catálogo de productos y de plataforma *Business to Business* (B2B) para sus distribuidores en un mismo sistema, lo que significaba que cubría todo el abanico de escenarios de relación con el visitante/cliente virtual (**de aquí en adelante visitante virtual por facilitar la comprensión al lector**). En este sentido, es necesario indicar que ORBEA S.Coop. es una empresa proveedora de 1º nivel, referente en la fabricación y distribución de bicicletas tanto a nivel nacional como internacional (Iturbe-Ormaetxe, 2018). Igualmente, es una organización con mucho interés en innovar. Esa innovación trae consigo el tener un mayor conocimiento, saber más sobre los visitantes virtuales de su sitio web. Dicho conocimiento podría ser beneficioso para su negocio. De hecho, en la investigación se tuvo en cuenta la relevancia del sector de la distribución a nivel internacional, tanto para las potencias internacionales contrastadas, como los Estados Unidos de América (USA) (IBM, 2012) como para los países emergentes como Brasil, Rusia, India, China y Sudáfrica (BRICS) (Indra, 2015). Además, también se tuvo en cuenta la relevancia a nivel europeo y

estatal. Para España y Europa suponía más de 2/3 del porcentaje total en importaciones y exportaciones, sin olvidar que era un 14 % de su Producto Interior Bruto (PIB) (Salcedo Barriga, 2012). Según la Asociación de Empresas del Gran Consumo (AECOC), el gran consumo representaba el 20 % del PIB estatal (Bonmatí, 2015). El sector de la distribución en España canalizaba en mayor o menor medida el desarrollo del comercio electrónico, a través del recorrido que estaban teniendo las plataformas de comercio electrónico. Dicho comercio electrónico suponía un punto de encuentro de clientes y no clientes, en definitiva visitantes virtuales, del cual resultaba fundamental comprender su comportamiento.

Para llevar este proyecto de investigación a cabo, fue necesario un proceso de comprensión que comenzaba con la revisión de la literatura, que analizaba desde el sector de la distribución y su cadena de suministro virtual, hasta el conocimiento del comportamiento del visitante virtual a través del seguimiento de objetivos, recogida, medición, reporte y análisis de los datos cuantitativos de Internet para optimizar los sitios web y las iniciativas de marketing web, *Web Analytics*. Todo ello, es la base a partir de la cual se plantearon las preguntas de investigación que fueron el soporte mediante el cual se construyó el modelo.

Por todo ello, y debido al *gap* que se identificó respecto al comportamiento de los visitantes virtuales en los proveedores (Castánder et al., 2017), y a las razones mencionadas anteriormente se decidió realizar el proyecto de investigación sobre ORBEA S.Coop.

1.4 Contribución

La contribución y/o novedad que ha aportado la investigación propuesta es la **generación de un modelo de comportamiento del visitante virtual en el marco de los proveedores de primer nivel en el sector de la distribución** para el caso de estudio de ORBEA S.Coop, ya que no se ha detectado ningún modelo en la revisión de la literatura sobre el comportamiento de los visitantes virtuales para los proveedores de primer nivel en el sector de distribución.

Capítulo 1. Introducción

Esta investigación tiene relevancia para dos ámbitos:

- Ámbito Científico-Académico: la elaboración de este proyecto de investigación permitirá disponer de un modelo, ya que no se ha identificado ningún modelo en la revisión de la literatura sobre el comportamiento de los visitantes virtuales para los proveedores de primer nivel en el sector de distribución, para cubrir el *gap*, sobre el comportamiento del visitante virtual en los proveedores de la cadena de suministro del sector de la distribución. Además, este nuevo modelo permitirá ahondar en la comprensión del comportamiento del visitante virtual aportando nuevas entidades, relaciones y variables.
- Ámbito Empresarial: es importante recalcar que este proyecto tiene como uno de los propósitos principales crear una nueva línea de producto que redundará en una diversificación de la oferta actual de i68. Incluyendo la mejora del ERP IZARO (producto de i68), ya que este proyecto va a ayudar a un crecimiento, desarrollando un nuevo producto que complementa y refuerza el impacto del nivel extendido del propio ERP. Además, también va a impactar en otros proyectos importantes de la unidad de I+D+i, al poder aplicar el conocimiento descubierto sobre los visitantes virtuales en el comercio electrónico. Asimismo, el resultado de este proyecto de investigación supondrá para los agentes del sector, una simplificación del aprendizaje sobre los visitantes virtuales, respecto a la utilización del comercio electrónico y su influencia en la gestión de las relaciones con los clientes. Siendo el punto de partida para el desarrollo de funcionalidades que permitan llevar a los sistemas de información interesados el modelo generado.

1.5 Preguntas de Investigación

Las preguntas de investigación soportan los objetivos específicos del proyecto y serán la base para construir el modelo. A continuación, se muestra a qué objetivo específico pertenece cada pregunta, las cuales, se explicarán más en detalle en el apartado de Metodología (véase capítulo 3).

1.5 Preguntas de Investigación

Tabla 1.1 Preguntas de Investigación por Objetivos Específicos

Objetivo Específico (OE) y Preguntas de Investigación (PI)

OE1: Conocer la evolución de la cadena de suministro desde la configuración física hasta la virtual.

PI.1.1 - ¿Cuáles son los componentes de la cadena de suministro y la relación entre ellos?

PI.1.2 - ¿Cuál es la evolución tanto de la cadena como del cliente físico al virtual?

OE2: Analizar el impacto que tiene en la propia organización, el comportamiento del visitante virtual en los escenarios de relación de la cadena de suministro.

PI.2.1 - ¿Cuáles son los diferentes escenarios de relación de ORBEA S.Coop. con los clientes/visitantes virtuales?

PI.2.2 - ¿Cuáles son las ramas de análisis de comportamiento del visitante virtual asociadas a cada escenario de relación?

PI.2.3 - ¿Qué conclusiones se pueden obtener de la observación del comportamiento de los visitantes virtuales de forma aislada y comparativamente con las ventas realizadas?

OE3: Identificar los factores que permiten aproximar a los proveedores de primer nivel, al resto de agentes de la cadena de suministro a nivel de comportamiento del visitante virtual.

PI.3.1 - ¿Cuáles de los valores de las características de las tecnologías de la información (IT) se tendrán en consideración en los escenarios de relación?

PI.3.2 - ¿Cuáles son las relaciones, tanto entre los tipos, como en el acceso a los productos visitados, en la web y los vendidos?

OE4: Definir la estrategia analítica para aplicar el conocimiento sobre el comportamiento del visitante virtual, en el desarrollo del sitio web y otros desarrollos tecnológicos.

PI.4.1 - ¿Cuáles son las pautas para construir una estrategia analítica?

Fuente: elaboración propia

Capítulo 1. Introducción

1.6 Metodología

En este proyecto de investigación la metodología (véase capítulo 3) ha seguido un enfoque cualitativo, ya que, a grandes rasgos, este tipo de enfoque propone investigar sobre el problema planteado mediante la generación de preguntas de investigación. Este tipo de enfoque sigue en la mayoría de los casos una lógica inductiva (Hernández Sampieri, Fernández Collado, & Baptista Lucio, 2006), es decir, el método científico es inductivo (Gray, 2013), alcanzando la generalización analítica planteando y respondiendo a preguntas de investigación.

Este método conlleva llevar a cabo el análisis de los datos obtenidos, con el propósito de identificar si existen patrones de comportamiento o de relaciones entre ellos. Partiendo de este punto, se pueden generar relaciones entre los elementos estudiados o incluso generar alguna teoría, evitando realizar inferencias basadas en los datos utilizados en el análisis.

1.7 Estructura de la investigación

Este proyecto de investigación ha constado de seis capítulos. En el capítulo 1, tal y como se ha descrito anteriormente, es la introducción general del trabajo, dónde se define: el propósito y los objetivos del proyecto; el planteamiento del problema; la contribución, las preguntas de investigación, la metodología, etc.

Mediante el capítulo 2 se ha realizado una revisión profunda del estado del arte que consta de los siguientes subapartados. En primer lugar, se trata el sector de la distribución, ya que este es el ámbito del proyecto investigación y tiene como objetivo revisar su evolución, ver sus características y determinar sus necesidades. A continuación, se entra de lleno en la cadena de suministro virtual, es decir, sus escenarios de relación y también las diferentes consideraciones respecto a las ramas de análisis de comportamiento. Finalmente, se analizan las técnicas de generación del conocimiento del visitante virtual mediante *Web Analytics*, ya que la finalidad de este apartado es ver qué fases hay que seguir para obtener dicho conocimiento sobre el comportamiento del visitante virtual y revisar las diferentes herramientas

1.8 Proyectos, publicaciones y transferencia a empresa

y/o técnicas en cada una de ellas.

A través del capítulo 3, se ha definido la metodología a llevar a cabo en el proyecto de investigación, que va a permitir recoger las evidencias cuyo análisis dará respuesta a las preguntas de investigación.

El análisis de la evidencia empírica se ha planteado en el capítulo 4, que tratará de responder a cada una de las preguntas de investigación.

Para el capítulo 5 se ha propuesto la generación del modelo de comportamiento del visitante virtual, para el caso de estudio de ORBEA S.Coop., con el propósito de facilitar al caso de estudio, la comprensión sobre los visitantes virtuales, a través de una estrategia analítica, tanto para valorar desarrollos tecnológicos futuros como para la parte comercial de ventas.

Por último, en el capítulo 6 se han descrito las conclusiones, limitaciones y líneas futuras de la investigación identificadas en esta tesis doctoral.

1.8 Proyectos, publicaciones y transferencia a empresa

A continuación, por un lado, se van a mostrar, brevemente, un conjunto de proyectos, publicaciones relacionadas con la elaboración de esta tesis doctoral y el registro de propiedad intelectual y transferencia de explotación empresarial realizada (ver Anexo), y por otro lado, otras publicaciones elaboradas durante la carrera profesional del doctorando.

1.8.1 Proyectos

- Proyecto: Web CAE (Año 2015)
 - Objetivo: Desarrollar una aplicación web para el *software* de gestión empresarial: Izaro ERP, para la coordinación de actividades empresariales

Capítulo 1. Introducción

(CAE) para facilitar y agilizar el intercambio de documentación con los proveedores.

- Proyecto: Web IRTEN (Año 2016)

- Objetivo: Desarrollar una aplicación web para facilitar el acceso y la gestión de incidencias de un CAU, para un entorno de mantenimiento industrial que puede ser beneficioso para el *software* de gestión empresarial: Izaro ERP.

Los dos proyectos siguientes han capacitado al doctorando en el conocimiento de la interoperabilidad entre el desarrollo web y la Base de datos, que ha sido una parte fundamental en el apartado de *Web Analytics* de la tesis.

- Proyecto: CRM Alerts (Año 2017)

- El objetivo de este proyecto fue obtener información a través de la consecución de datos web y que pueden ser beneficiosos para el *software* CRM de Izaro ERP: Izaro CRM. Este proyecto aportó el conocimiento necesario para el funcionamiento de técnicas que obtienen datos web, para luego comprender mejor la operativa de los procedimientos que utiliza el proceso de *Web Analytics*.

- Proyecto: Plataforma Comercial (Año 2018)

- El propósito de este proyecto fue desarrollar una plataforma, que a través de una aplicación web aunase los diferentes sistemas de comercio electrónico que se ofrecen en, el *software* de gestión empresarial, Izaro ERP. Este proyecto aportó el conocimiento necesario para la comprensión tanto conceptual como tecnológica del funcionamiento de los diferentes escenarios de relación (B2B, B2C) sobre los que se ha realizado el proyecto de tesis.

- Proyecto: Izaro Analytics (Año 2018)

- El objetivo de este proyecto fue desarrollar un cuadro de mando que sirviese para analizar los datos provenientes de los *log* de uso del *software* de gestión empresarial, Izaro ERP. Este proyecto aportó el conocimiento del

1.8 Proyectos, publicaciones y transferencia a empresa

flujo que va desde el origen de datos del proyecto hasta el cuadro de mando, así como el desarrollo y la utilización del propio cuadro, necesario para realizar los posteriores análisis de datos.

1.8.2 Publicaciones

A continuación, en primer lugar, se muestran todas las publicaciones realizadas durante la formación de la tesis doctoral y en segundo lugar, un conjunto de artículos desarrollados durante la trayectoria profesional del doctorando que le ayudaron al propio doctorando a ampliar sus conocimientos.

- Título: Metodología para el comportamiento de los visitantes virtuales en los proveedores del sector de la distribución.

Castánder, I., Serna, A., & Busto, I. (2017). Metodología para el comportamiento de los visitantes virtuales en los proveedores del sector de la distribución. *Distribución y consumo*, 27(149), 66.

El propósito de este artículo es desarrollar una metodología para llevar a cabo un modelo de comportamiento del visitante virtual en el marco del sector de la distribución. De hecho, se facilita una vía para posteriores investigaciones sobre la cadena del sector desde la perspectiva física a la virtual, gracias a la exhaustiva revisión de la literatura llevada a cabo y a la identificación de las áreas de conocimiento que abarcan. Partiendo de dicho análisis, se establece el foco en las áreas de conocimiento sobre el comportamiento de los visitantes virtuales en la parte de la cadena del sector donde se ha identificado un *gap* con una problemática que justifica el desarrollo de la propuesta metodológica.

- Título: Estudio comparativo de las herramientas de business intelligence: Empoderando el criterio de selección a las pymes.

Castánder, I., Gerrikagoitia, J. K., & Alzua-Sorzabal, A. (2013, dic). Estudio comparativo de las herramientas de business intelligence: Empoderando el criterio de selección a las pymes. *tourGUNE Journal of Tourism and Human Mobility*, Issue 1 , 45-56.

Capítulo 1. Introducción

Aportación para el conocimiento para la tesis: Herramientas de visualización y análisis de datos.

- Título: New trends of intelligent e-marketing based on web mining for eshops.

Gerrikagoitia, J. K., Castánder, I., Rebón, F., & Alzua-Sorzabal, A. (2015, feb). New trends of intelligent e-marketing based on web mining for eshops. Procedia - Social and Behavioral Sciences, 175 (0), 75-83.

Aportación para el conocimiento para la tesis: Tratamiento con diferentes técnicas de análisis de datos.

- Título: Sistema regional de información turística para euskadi integración de distintas fuentes y nuevos indicadores para una visión holística del turismo.

Castánder, I., Zurutuza, M., Gerrikagoitia, J. K., & Alzua-Sorzabal, A. (2013, dic). Sistema regional de información turística para euskadi integración de distintas fuentes y nuevos indicadores para una visión holística del turismo. tourGUNE Journal of Tourism and Human Mobility, Issue 1 , 35-43.

Aportación para el conocimiento para la tesis: Tratamiento con diferentes y diversas fuentes de datos de un origen de datos tecnológico.

- Título: An antifraud system for tourism smes in the context of electronic operations with credit cards.

Rebon, F., Castánder, I., Argandoña, J., Gerrikagoitia, J. K., & Alzua-Sorzabal, A. (2015, feb). An antifraud system for tourism smes in the context of electronic operations with credit cards. American Journal of Intelligent Systems, 5 (1), 27-33.

Aportación para el conocimiento para la tesis: Tratamiento con diferentes técnicas de análisis de datos.

El tiempo es importante, transcurre y no se puede volver atrás.

Anónimo

CAPÍTULO

2

Estado del Arte

En esta sección, se va a describir el estado del arte del proyecto de investigación, partiendo del sector de la distribución y su cadena de suministro virtual, pasando por las comparativas de herramientas de *web tracking*, hasta el conocimiento del comportamiento del visitante virtual. Dónde los modelos estaban sobre todo orientados al consumidor final, etc. Esto supone una falta de investigación sobre otros participantes de la cadena, importantes todos ellos en el recorrido realizado, como es el caso de los proveedores.

2.1 El sector de la distribución

La distribución es el sector que engloba a las organizaciones orientadas a la concentración masiva de productos o servicios a un amplio espectro de consumidores. Es el sector que entrega productos al cliente o al consumidor final. Este sector es muy importante en países principales como USA, donde el consumo agregado representa el 70 % del PIB (IBM, 2012), y también en los BRICS, debido sobre todo al peso en el crecimiento que tienen y que van a tener estos países: según el Fondo Monetario Internacional (FMI) el 70 %. En este sentido, un dato a tener en cuenta es que tanto India como China conformarán el 40 % del porcentaje anterior (Indra, 2015).

Capítulo 2. Estado del Arte

2.1.1 Estructura

El sector de la distribución ha tenido diversas estructuras a lo largo de historia. Comenzando de por los grandes almacenes de las ciudades importantes, pasando por las tiendas descuento y grandes superficies orientadas a los núcleos de población, hasta la irrupción del comercio electrónico. En este sentido, dentro de la evolución de la estructura del sector, la cadena de suministro juega un papel fundamental, debido al rol que tiene la distribución en la propia cadena. Por ello, es necesario saber qué es exactamente la cadena de suministro y la gestión de la misma. Según Busto (2012) la cadena de suministro es "la unión de todos los agentes que hacen posible que un producto salga al mercado, o lo que es lo mismo, la unión de todas las empresas que participan en el abastecimiento, producción, almacenamiento, comercialización y distribución de un producto y sus componentes."

Asimismo, la gestión de la cadena de suministro o *Supply Chain Management* (SCM) según el *Council of Supply Chain Management Professionals* (CSCMP) (of Supply Chain Management Professionals (CSCMP), 2018) se refiere a "la planificación y gestión de todas las actividades involucradas en la distribución y abastecimiento, en la transformación y la gestión logística. También incluye la coordinación y colaboración con los socios de la cadena que pueden ser proveedores, intermediarios, proveedores de servicios y clientes. En esencia, la gestión de la cadena de suministro integra la gestión del suministro y la demanda dentro y entre compañías (p.1.)".

Una vez descritos ambos conceptos es evidente lo importante que es la distribución para que los productos lleguen al cliente final y se complete el suministro. Y por supuesto llevar a cabo una gestión de las actividades y demás elementos que van a llevar a cabo dicha distribución a través del SCM. A todo esto hay que añadir que las cadenas de suministro también han tenido su propio desarrollo existiendo diferentes categorías asociadas a los siguientes sistemas de producción: *push*, *pull* y *push-pull*.

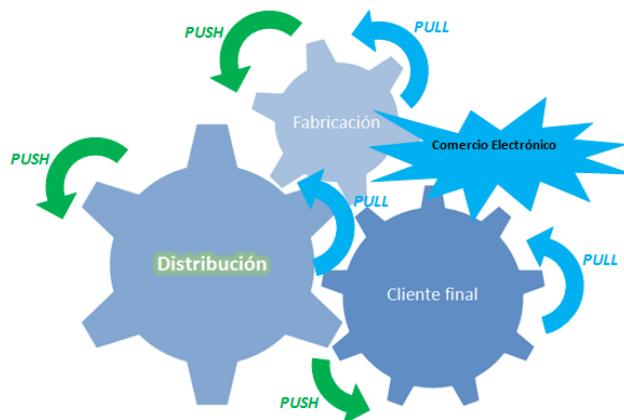
- **Push:** en este tipo de sistema, la cadena de suministro empuja hacia la demanda. Es decir, a partir de estimaciones sobre la demanda se generan las

2.1 El sector de la distribución

órdenes de producción.

- **Pull:** en este otro tipo, la cadena de suministro es arrastrada por las peticiones de la demanda. De esta forma, se desencadena un proceso donde cada elemento de la cadena demandará a su anterior, según consuma el material requerido (Busto, 2012).
- **Push-pull:** este es el sistema híbrido, el cual combinaría parte de la cadena con uno de los sistemas (por ejemplo *push*) y otra parte con el segundo tipo (*pull*). Este sistema ha ido evolucionando en los últimos años respecto a las cadenas de valor (Porter & Millar, 1985). Esta evolución ha permitido la desestructuración de las mismas e integrando la actividad de comercio electrónico como un elemento importante para la parte *pull* del sistema reforzando de esta manera al bloque del distribuidor (ver Figura 2.1).

Figura 2.1: Modelo Push-Pull en la cadena de valor



Fuente: Elaboración propia basado en (Indra, 2015)

Como se puede observar en la figura anterior, en la parte de distribución el comercio electrónico es importante. Por lo tanto, este fenómeno también se puede trasladar a la cadena de suministro sobre todo a la asociada a una estructura del sector más digital o virtual. Por todo ello, se va a realizar una comparativa de lo que ha venido siendo la cadena de suministro tradicional o física y la cadena virtual que se está gestando. No obstante, antes de ello a continuación se describen las características y necesidades del sector.

Capítulo 2. Estado del Arte

2.1.2 Características y necesidades del sector

Una vez vista la estructura del sector, en este subapartado se describen las características y necesidades del mismo en base a un conjunto de cambios importantes identificados para la distribución desde un nivel internacional hasta un nivel estatal.

Tabla 2.1 Características del sector de la distribución estatal

Cambio	Característica
Cuatro gigantes están transformando el panorama de la distribución.	Monopolio
Los países emergentes ofrecen oportunidades de crecimiento.	Crecimiento
Los hijos de la llamada "generación del milenio" habrán tomado el control.	Control generacional
Los consumidores seguirán siendo "bipolares".	La bipolarización
Demasiado espacio para tan pocas ventas.	Sobre-dimensión
El comercio electrónico plantea retos para el modelo de negocio de la mayoría de distribuidores con tiendas físicas.	Consciencia transformadora
Las plataformas móviles, los sensores inteligentes y las redes sociales fomentan nuevas formas de compra.	Convivencia tecnológica

Fuente: elaboración propia basada en (IBM, 2012; Indra, 2015)

Como se puede observar en la tabla anterior, en el sector de la distribución un conjunto de cambios están generando unas características que componen un recorrido más que interesante: empresas, canal y usuarios.

Respecto a las particularidades más propias del lado de las **empresas**, el monopolio generado por los gigantes del sector se compagina con la posibilidad de crecimiento que tiene el sector en los países emergentes.

Relacionadas con las anteriores están las características de **canal**. En este sentido,

2.1 El sector de la distribución

la sobre-dimensión existente en las superficies por la falta de ventas es un problema que gracias a la conciencia transformadora que se está gestando en el sector, permite que el comercio electrónico sea una vía de solución. De hecho, estas soluciones conllevan estar conviviendo con tecnología de múltiples ámbitos como redes sociales, plataformas móviles, sensores inteligentes, etc.

Respecto a las características relacionadas con los **usuarios**, a medida que se avance en el control generacional, por parte de los hijos de la llamada "generación del milenio", la convivencia tecnológica y la conciencia transformadora, comentadas anteriormente, tendrán un impulso mayor que el actual. Aun así, esto no significa que los consumidores dejen de estar en una situación de bipolarización, es decir, que a pesar de la pérdida de poder adquisitivo se mantengan las ventas en productos de marcas de lujo, *discounters*, etc.

A continuación, se describen un conjunto de necesidades que el sector de la distribución tiene que cubrir. La identificación de las mismas se ha podido realizar gracias a la revisión de diversos estudios (AECOC, 2009a, 2009b; IBM, 2012; Salcedo Barriga, 2012; Guerrero, 2015; Indra, 2015).

Las necesidades identificadas son las siguientes:

1. **Aumentar el volumen de ventas**, la crisis ha reducido la cantidad de dinero que los consumidores están dispuestos a gastarse en ellos mismos y en los demás. Se ha convertido en la **"nueva normalidad"**. Un concepto que el Banco Mundial también tiene en consideración y en el cual se da la situación de un crecimiento moderado o inexistente. A pesar de ello, el sector de la distribución ya ha vivido antes drásticas transformaciones. Por ello, son cuatro los factores a tener en cuenta: la situación económica, los cambios demográficos, las innovaciones tecnológicas y el comportamiento de los clientes. Este último, es fundamental debido a que conociendo su comportamiento se dispone de una mayor comprensión a cerca de cuándo, cómo y dónde pueden llevar a cabo la compra. Además, si a lo comentado se le añade el punto de dinamismo cada vez más frecuente, el conocimiento del comportamiento adquiere un grado mayor en importancia

Capítulo 2. Estado del Arte

para el sector de la distribución.

Por todo ello, es necesario aumentar el volumen de ventas y no confundirlo con otro concepto directamente relacionado como es el número de las ventas. Es decir, el primero es la cantidad total de euros y el segundo son las transacciones realizadas.

2. **Superar las grandes barreras de entrada al sector**, desde el punto de vista del consumidor que existan grandes barreras de entrada al sector de la distribución es negativo. Es decir, que empresas del sector no puedan acceder a un mercado debido a las exigencias de la legislación del país en cuestión, a que la oferta principalmente incluye *comodities* (productos con poco valor añadido), o a la trascendencia de las economías de escala.

Todo ello, genera que solamente las multinacionales del sector puedan saltarse dichas barreras y de esta forma, apropiarse de la mayor cuota de mercado posible. Un mercado en el cual pueden disponer de una actitud dominante respecto al consumidor, ya que este último queda indefenso ante el dominio de estas organizaciones. Además, con la actual coyuntura económica el poderío será mayor si se sigue limitando la inversión en innovación, ya que una de las vías para paliar esta necesidad está relacionada con otra de las necesidades del sector de la distribución como es la de proporcionar productos de valor añadido.

3. **Facilitar el control de productos**, en el sector de la distribución otro punto a tener en cuenta es la dificultad con la que el fabricante se encuentra a la hora de conocer el estado del producto generado durante el recorrido hasta la venta al consumidor final. En este sentido, la situación del producto puede variar según en manos de quién esté y si permite al proveedor acceder o conocer el estado del producto. Aun así, existe un caso en el cual se pierde el control del producto debido a la pérdida del mismo. Este fenómeno se conoce como "*shrinkage* o pérdida desconocida". Según la AECOC (2009a) la pérdida desconocida es "la diferencia contable entre un stock teórico y un stock físico". La cual genera, según AECOC "pérdidas de más de 1.700 millones de euros al año, lo que equivale al 1 % de la facturación de las empresas españolas de distribución y

2.1 El sector de la distribución

que repercute en toda la cadena y al consumidor”. Por ello, AECOC a través de su informe sobre recomendaciones sobre la prevención de la pérdida desconocida (AECOC, 2009b), lleva desde hace unos años queriendo concienciar a las empresas del sector para que realicen un mayor control de los productos adquiridos a través de la aplicación de determinadas buenas prácticas para disminuir la pérdida desconocida.

En este sentido, Guerrero (2015) afirma:

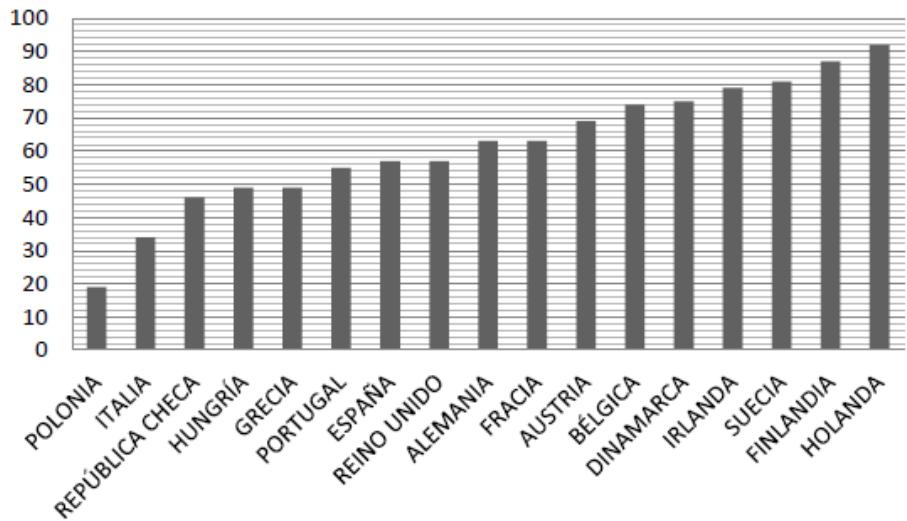
la pérdida desconocida en los comercios españoles se sitúa en el índice más bajo desde 2010, según el Barómetro Mundial del Hurto en la Distribución. Entre 2001 y 2007, el hurto descendió pasando de un 1,47 % sobre el total de ventas al 1,28 %. El inicio de la crisis económica mundial supuso un repunte de la pérdida desconocida alcanzando el nivel más elevado en 2011 (1,40 %) (p.1).

Aun observando que los datos van en la dirección correcta, hay que seguir disminuyendo la pérdida desconocida y facilitar el control de productos. Por ello, aparte de las acciones que se llevan a cabo en las grandes superficies es necesario que el comercio electrónico tenga cada vez más presencia en el sector.

4. **Controlar/Disminuir la concentración actual del sector**, la concentración del sector de la distribución en España, todavía no alcanza los niveles de determinados países europeos (ver Figura 2.2). Aunque está en una situación creciente y hay que tenerla en cuenta, ya que las principales empresas del sector cada vez tienen más control sobre el mercado debido a sus porcentajes de cuota.

Capítulo 2. Estado del Arte

Figura 2.2: Porcentaje de cuota de mercado de las cinco primeras empresas según el país europeo



Fuente: (Salcedo Barriga, 2012)

Para cumplir con esta necesidad, en primer lugar debería desarrollarse una ley que fuese en favor de las Pequeñas y Medianas Empresas (PYMES) y así éstas puedan disponer de un mayor espacio en el mercado entre las grandes organizaciones del sector. En segundo lugar, debido a que en España el consumidor final sigue prefiriendo realizar compras en las superficies (tiendas, supermercados, etc.), se tiene que fomentar la implementación del comercio electrónico en las PYMES del sector (Indra, 2015). De este modo, se favorece la necesidad de conocer al visitante virtual.

5. **Proporcionar al mercado productos de valor añadido**, es el factor sorpresa sobre el consumidor para marcar diferencias con los competidores. Esta necesidad a nivel estatal tiene su origen en la existencia de una alta competencia en la cual la variable principal o la única, es el precio del producto. Esta situación está generando una batalla que puede dañar seriamente al sector de la distribución. Por ende, para poder sorprender al consumidor y evitar entrar y/o permanecer en una guerra de precios es necesario proporcionar al mercado productos de valor añadido. Productos que sorprendan, productos que no dependan de ser adquiridos meramente si su precio es más barato que el de la competencia.

2.1 El sector de la distribución

Asimismo, es necesario acompañar el lanzamiento de productos de valor añadido con actividades que controlen la pérdida de clientes (*Customer Churn*). Es decir, es fundamental para las empresas del sector que la pérdida de clientes sea controlada y para ello es necesario predecir cuales de ellos van a dejar de serlo. Sobre todo porque para que el control de la pérdida de clientes sea efectivo a posteriori, es necesario realizar campañas de retención y si no es posible predecir dicha pérdida, el coste de las campañas de retención de clientes es mayor. Este paradigma dentro del comercio electrónico, si se tiene en consideración el valor medio de los clientes es más crítico en el B2B que en el B2C (Tamaddoni Jahromi, Stakhovych, & Ewing, 2014).

En cambio, en términos generales es más fácil que se produzca infidelidad en el B2C que en el B2B, aunque sea por el mero hecho de que en el segundo existe una relación contractual.

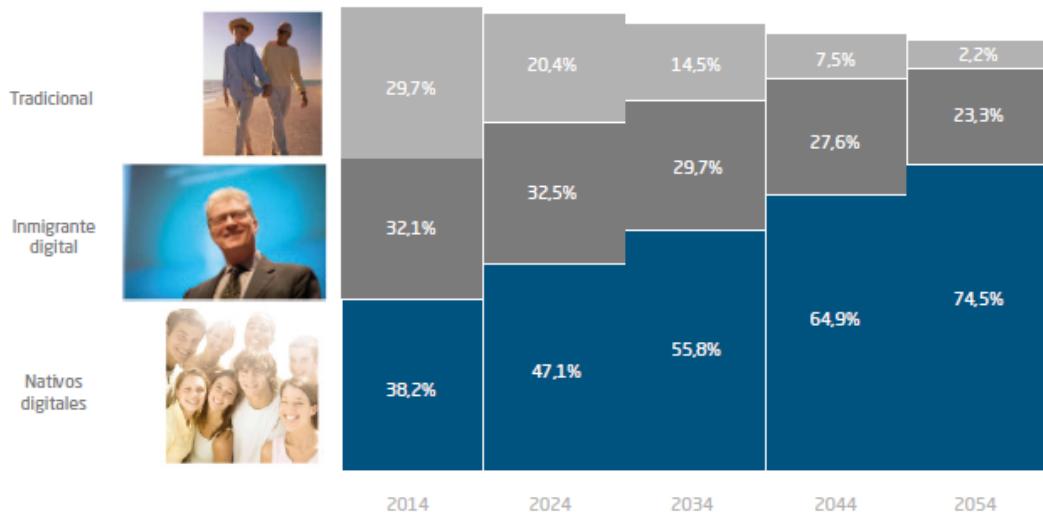
6. **Adaptarse a los cambios de demanda**, menos tiempo de compra, mas acceso a la información. El ritmo frenético (sin tiempo, nerviosismo,...) existente en la sociedad, proporciona una oportunidad a las organizaciones para que puedan proporcionar productos o servicios que permitan ganar tiempo, reducir nerviosismo, etc. a sus clientes. En resumen, unos servicios que tienen que amoldarse a la demanda. Es en este contexto dónde debido a estos cambios los visitantes (consumidores) virtuales pueden adquirir diversas tipologías o roles. Por ejemplo, los que están preocupados por el entorno medioambiental y la ecología, o aquellos que se convierten ellos mismos en los propios fabricantes, etc. Este último caso, se debe a que los visitantes (consumidores) son cada vez más domésticos. Lo cual ha propiciado a que muchos de ellos hayan pasado de ser netamente consumidores a convertirse en fabricantes de productos (Indra, 2015). Esto posibilita, que los visitantes virtuales también puedan ser fabricantes o distribuidores.
7. **Convertir en visitantes (consumidores) virtuales a personas de avanzada edad**, que no tengan conocimientos de Internet y tengan problemas de desplazamiento. Con el paso del tiempo, la ciudadanía y el avance de la tecnología han

Capítulo 2. Estado del Arte

ido perfilando diferentes roles o tipos visitantes (consumidores) en la sociedad. Según Indra (2015) el visitante del futuro pasa por los siguientes tipos: tradicionales, inmigrantes digitales y nativos digitales. A nivel estatal, en el 2015 los nativos digitales representaban el 38 % de la población, que aumentaría al 56 % en 20 años y al 75 % en 40 años. Asimismo, con el paso de los años y con la aparición de nuevas generaciones de nativos digitales (ver Figura 2.3), este segmento de población vivirá en una sociedad totalmente digitalizada y con conocimientos tecnológicos adquiridos cada vez con edad más prematura. Este hecho, supone un paso importante respecto a los inmigrantes digitales, que tenían sus primeros contactos con la tecnología a edades más avanzadas e incluso cercanas a la adolescencia o a la mayoría de edad.

Según Marc Prensky, quién acuñó el término “nativo digital”, para 2020 las personas en todo el mundo estarán conectadas a la AORTA (*Always On Real-Time Access*).

Figura 2.3: Evolución del peso de los segmentos de consumidor (sobre la población total)



Fuente: (Indra, 2015)

8. **Enfocar el futuro del canal tradicional dentro del sector**, los profesionales del sector consideran que este canal se va ir reduciendo con el paso del tiempo

2.1 El sector de la distribución

y que aun así se tienen que tomar iniciativas mientras perdure. Estas iniciativas incluyen el optimizar superficies, generar tipologías y fidelizar al cliente, etc. No obstante, a día de hoy para un sector de la población consumir a través del canal tradicional, supone una razón para socializar y pasar el tiempo de ocio.

Una vez vistas las características y necesidades del sector de distribución, a continuación se describe la comparativa entre las cadenas de suministro física y virtual.

2.1.3 Comparativa entre las cadenas de suministro física y virtual

En la actualidad, la cadena de suministro física que se ha conocido hasta ahora, va a ir perdiendo protagonismo en favor de la cadena de suministro virtual (*Virtual Supply Chain* o *e-Supply Chain*). Esta última, según Manthou et al (2004) es "las comunicaciones y las operaciones de la columna vertebral de una red virtual, que enlaza a partners de negocio, proveedores y clientes juntos como una entidad cohesionada y colaborativa."

Así, hay que añadir que el uso de las TIC, ya no es solo para unir a las partes sino que el propio intercambio de datos entre determinadas partes de la cadena, genera la creación de la misma. Las cadenas de suministros virtuales se basan más en información que en inventario (Christopher, 2000).

Una información que junto con los otros flujos como el financiero y el transporte siguen respetando los procesos existentes en la cadena física.

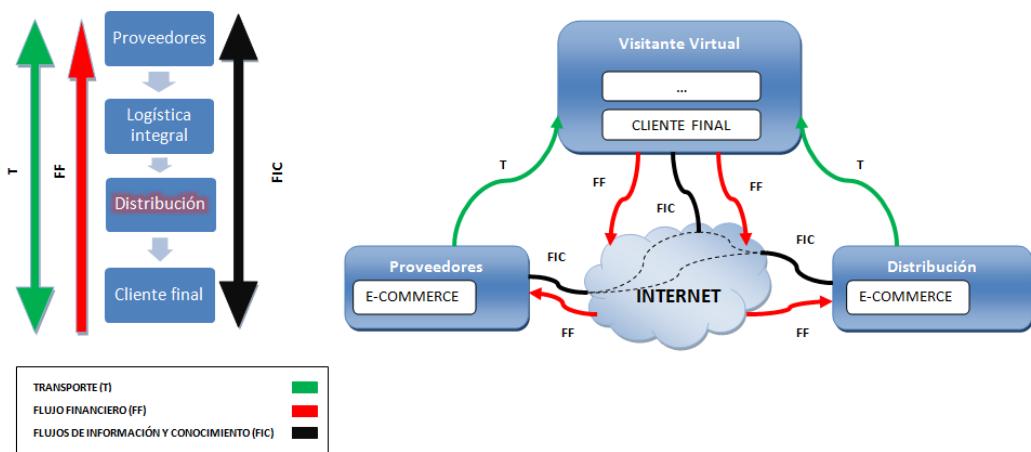
En este sentido, es importante recordar lo mencionado por Busto (2012):

la cadena de suministro es algo más que la logística y abarca, además de los procesos logísticos (flujos de materiales), la gestión de los flujos de información, los flujos financieros e incluso, el capital intelectual (conocimiento) asociado a todos ellos (p.21).

Capítulo 2. Estado del Arte

A continuación, en la Figura 2.4 se pueden observar las dos cadenas y cómo a pesar de la influencia de Internet en los flujos de información y conocimiento y en los flujos financieros, en lo que es el transporte todavía es necesaria la capa física. En definitiva, los flujos se mantienen pero su rol o su peso varían según la cadena de suministro. Además, desde la perspectiva de cliente en la cadena de suministro virtual, está entidad estará encapsulada dentro del visitante virtual, lo cual genera la necesidad de conocer al mismo. Sobre todo, si se tiene en cuenta que no conocer a los clientes aumentará la marcha de los mismos en las cadenas de suministros virtuales que se basan en pedidos únicos y una vez realizados los clientes desaparecen de la misma (Chandrashekhar & Schary, 1999).

Figura 2.4: Comparativa entre las cadenas de suministro física y virtual



Fuente: Elaboración propia basado en (Pant, Sethi, & Bhandari, 2003; Manthou et al., 2004; Busto, 2012)

Una vez vista la comparativa, es necesario conocer el recorrido existente que va desde la cadena física hasta la cadena virtual. Para ello, en el siguiente punto se describe el recorrido en base a los componentes de la cadena y una retrospectiva.

2.1.4 Recorrido de los componentes de la cadena: desde la perspectiva física a la virtual

Los diferentes participantes de la cadena de suministro han vivido una evolución en lo referente a sus relaciones, intercambio de datos, información, impacto de las

2.1 El sector de la distribución

TIC, resultados y relación en base a la marca, etc. En este sentido, a nivel de la cadena de suministro física, hay que destacar el trabajo que realizaron Yagüe & Rubio (2006) sobre las relaciones de los fabricantes y distribuidores de la cadena. Este estudio realizaba una investigación empírica según los rendimientos y vínculos de las marcas de los mismos, teniendo en cuenta diferentes variables de negocio (margen bruto, ventas, etc.) y otras como satisfacción y dependencia. Tras aplicar un análisis estadístico a los datos provenientes de encuestas, una de las conclusiones principales fue que "la investigación corrobora que el fabricante obtiene un margen de beneficios inferior con la marca del distribuidor (Quelch y Harding, 1996) (p. 83)".

Esta investigación, ya indicaba un salto a nivel de resultados y relaciones en base a la marca entre los proveedores de la cadena de suministro. La diferencia si fuera en aumento no sería beneficioso, ya que ambos tienen que satisfacer las necesidades del consumidor, tal y como se comentaba en el propio estudio.

A pesar de ello, hay que recalcar que la evolución de los resultados y relaciones en base a la marca entre los participantes de la cadena de suministro, no está ligado directamente con el recorrido realizado por la cadena de suministro hacia su versión virtual. Sobre todo, a nivel de crecimiento en marcas de fabricantes, de distribuidores y segmentos de mercado de las propias marcas. En este sentido, otro trabajo de Rubio & Yagüe (2009) fue más allá indagando en los segmentos disponibles en el mercado de las marcas de las tiendas, para tratar las diferencias entre las categorías existentes.

Continuando con otros aspectos, aparte de las relaciones y la marca, existe bastante literatura respecto al impacto de las TIC en la cadena de valor de suministro y respecto a diferentes impactos como pueden ser beneficios, reducción de costes, etc. De hecho, hay que recalcar estudios como el que propusieron Barbolla, Román & Yagüe (2013), en el cual se desarrollaba un modelo de antecedentes de beneficios relacionales en el canal de distribución, en base a mediciones en torno a conceptos como confianza, satisfacción y compromiso. Concluía indicando que se tenían que mejorar las relaciones e intercambio de información entre los componentes y sobre todo de cara al consumidor final. Esto viene a constatar que en las cadenas de suministro, e incluso en las denominadas BOSC (*Build-to-Order Supply Chain*) se centraban en los consumidores más que en otros agentes de la cadena

Capítulo 2. Estado del Arte

(Gunasekaran & Ngai, 2005) y que en las denominadas cadenas de suministro sostenibles las métricas tenidas en cuenta responden a conceptos muy genéricos como económicas, entorno y sociedad que dan pie a indicadores de negocio (nº de contratos, pedidos de compra, etc.) (Hassini, Surti, & Searcy, 2012).

Respecto a la relación entre proveedores y clientes existe suficiente literatura en relación a características como puede ser la confianza, la gestión del riesgo en el suministro, etc. En el caso de la confianza, el trabajo realizado por Liu, Tao, Li & El-Ansary (2007) tomaba el testigo en lo que concierne a construir un modelo basado en tipos de confianza (honestidad, benevolencia, directa e indirecta). De hecho, los autores planteaban previamente un conjunto de hipótesis que son refutadas mediante la aplicación de estadística a datos obtenidos de una muestra de distribuidores, siendo el modelo el que recogía dichos resultados mostrando el comportamiento de los tipos de confianza en la relación proveedor y distribuidor. Del modelo se concluye que existen dos vertientes que ejecutan este tipo de relaciones: 1) el acuerdo contractual y 2) la propia naturaleza de la relación. La primera es importante sobre todo en contextos de sistemas de comercio electrónico de tipo B2B y la segunda, lo es también, ya que la relación en sí es un factor clave en la cadena de valor del sector de la distribución. En este sentido, no hay que olvidar que las relaciones más cercanas son las más propensas a romperse (Anderson & Jap, 2005).

En el caso de la gestión del riesgo, hay que recalcar los estudios realizados por Wu & Olson (2008) y Tang & Tomlin (2008) sobre los riesgos en la cadena de suministro y como mitigar su impacto respectivamente, junto con el estudio realizado por Kull & Closs (2008), mediante el cual se evaluó la probabilidad y magnitud de los riesgos en la cadena ante diferentes contextos. Este trabajo, demuestra la correlación entre los fallos en el suministro con determinadas situaciones, como por ejemplo, el fallo en los proveedores de segundo nivel. Además, se propuso una línea a investigar sobre la gestión de riesgos de la cadena en el contexto de los proveedores de primer nivel.

A pesar de cualidades como la confianza o la gestión del riesgo en la cadena, los vínculos entre los componentes de la cadena fue transformándose a medida que Internet comenzó a ser un enlace electrónico para los componentes de la cadena, permitiendo la extensión de los modelos de negocio e iniciándose la gestación de

2.1 El sector de la distribución

la cadena de distribución virtual. A través de esta última, también se desarrollaron los sistemas de comercio electrónico, punto de encuentro entre clientes, no clientes,... en definitiva visitantes virtuales.

Esto dio pie a estudios que indagaron en el intercambio de datos como opción a integrar a los componentes de la cadena. En este sentido, las investigaciones que realizaron, por un lado, García-Dastugue & Lambert (2003) y por otro lado, R. Lancioni, Schau & Smith (2003) trataron la generación de un *framework* o marco de trabajo sobre los mecanismos de Internet y flujos de coordinación en contextos de B2B y B2C. Recalcando la importancia de los *information hubs/flows* (Bal, Wilding, & Gundry, 1999) o puntos de intercambio de información orientados sobre todo a los diferentes niveles de proveedores y a los sistemas de comercio electrónico. Sin olvidar, que se mencionan indicadores muy cercanos como por ejemplo el volumen de transacciones, el cual está muy relacionado con el comportamiento del visitante virtual.

Igualmente, otros autores como Lee & Wang (2004) investigaron el intercambio de datos desde el punto de vista del impacto del *e-business* en la cadena de suministro. De hecho, en este trabajo se ponían en valor nuevos modelos de negocio que remarcaban el concepto de *cross docking*, mediante el cual los diferentes componentes de la cadena podían compartir información sobre el catálogo de sus productos/servicios u otro tipo de información con los otros participantes a través del *information hub*, facilitando de esta manera la integración de los miembros de la cadena.

No obstante, a medida que las TIC e Internet iban evolucionando, los estudios respecto los visitantes virtuales se fueron orientando hacia conocer la aceptación de la tecnología bajo diferentes técnicas como TAM (*Technology Acceptance Model*), IDT (*Innovation Diffusion Theory*), etc. (Gillenson, Sherrell, et al., 2002; L.-D. Chen & Tan, 2004), con el propósito de aterrizar para el sector la incipiente irrupción de los sistemas de comercio electrónico. Tales sistemas, como pueden ser las plataformas de B2B y las tiendas online (B2C) han permitido flexibilizar la operativa de la cadena generando flexibilidad y mejor acceso para los agentes que la componen. Por ello, se logró el objetivo de integrar la cadena de suministro

Capítulo 2. Estado del Arte

con los sitios web que disponían de sistemas de comercio electrónico, mediante el desarrollo de arquitecturas y plataformas integradoras (Ghenniwa, Dang, Huhns, & Shen, 2006). Incluso se diseñaron aplicaciones web para que los proveedores gestionasen su relación dentro de la cadena de suministro con los vendedores (James, Grosvenor, & Prickett, 2004).

Sin embargo, los modelos que estudiaban la aceptación de la tecnología querían facilitar que las organizaciones se decantasen cada vez más por estos tipos de canal de venta, siendo estos modelos a su vez basados en encuestas, observaciones, etc. Es decir, los datos que recibían de los visitantes virtuales no eran los datos de comportamiento que se generaban mediante la utilización de los canales de venta, sino que eran respuestas a preguntas predefinidas e incluso verificaciones de hipótesis planteadas directamente. De hecho, hay que volver a recalcar que estos modelos ponían el foco en la aceptación de la tecnología y no en la navegación en base a los movimientos que pueda realizar el visitante virtual a través del canal de venta que estuviesen incorporados al respectivo sitio web, suponiendo un vacío en lo referente al comportamiento de los visitantes virtuales en la cadena de suministro.

De hecho, estos modelos disponían de su muestra de encuestas, observaciones,... en base a registros generados por personas relacionadas a los diferentes componentes de la cadena de valor del sector de la distribución, y sobre todo en concreto al consumidor final. **Esto supone una laguna en la investigación sobre otros participantes de la cadena, importantes todos ellos en el recorrido realizado, como es el caso de los proveedores.**

Por último, en la Tabla 2.2 se muestra una revisión de la aportación bibliográfica que se ha realizado en este apartado.

2.1 El sector de la distribución

Tabla 2.2 Recorrido desde la parte física hasta la virtual en la cadena de suministro del sector de la distribución.

Estudio	Autor
Distribuidores y fabricantes: Rendimientos y evolución de la marca	Yagüe and Rubio (2006)
Categorizaciones en el mercado de las marcas	Rubio and Yagüe (2009)
Estudio de conceptos como confianza, satisfacción y compromiso	Barbolla et al. (2013)
BOSC (Build-to-Order Supply Chain)	Gunasekaran and Ngai (2005)
Métricas e indicadores de negocio en las cadenas de suministro sostenibles	Hassini et al. (2012)
Característica de la relación: la confianza	Y. Liu et al. (2007)
Cercanía en la relación	Anderson and Jap (2005)
Característica de la relación: la gestión del riesgo	Wu and Olson (2008), Tang and Tomlin (2008), Kull and Closs (2008)
Mecanismos de Internet y flujos de coordinación en contextos B2B y B2C	(Bal et al., 1999) , García-Dastugue and Lambert (2003), R. Lancioni et al. (2003)
Impacto del <i>e-business</i>	Lee and Whang (2004)
Aceptación de la tecnología	Gillenson et al. (2002), L.-D. Chen and Tan (2004)
Desarrollo de arquitecturas y plataformas integradoras	(Ghenniwa et al., 2006), (James et al., 2004)

Fuente: Elaboración propia

2.1.5 Retrospectiva: de la integración al comportamiento en los canales de venta

En el apartado anterior, se ha visto el recorrido realizado por los componentes de la cadena de distribución hacia el entorno virtual y como la orientación hacia los consumidores finales y la falta de estudio del comportamiento del visitante virtual sobre otros participantes de la cadena, presentan una falta de investigación al respecto.

Por ello, y para poder indagar más en dicho *gap*, se ve necesario realizar una retrospectiva en el tiempo y tratar más en profundidad otros estudios y sobre todo establecerse en los contextos de B2B y de B2C.

Capítulo 2. Estado del Arte

A finales del siglo pasado y principios del actual, se realizaron trabajos que trataban la integración desde una perspectiva estratégica (Johnson, 1999; Christopher & Jüttner, 2000). Otros desde una perspectiva tecnológica orientados al Marketing (MIS) (Trappey & Trappey, 1998) y de gestión de relaciones (Lambert & Cooper, 2000)). Además, de investigaciones sobre la competencia y la concentración en los canales de distribución (Roche et al., 1999) y otros, en los que se investigaban el impacto del comercio electrónico sobre las relaciones de la cadena de suministro. Dentro de este último bloque, hay que destacar el trabajo realizado por Golicic, Davis, McCarthy & Mentzer (2002), en el cual se investigaba el impacto sobre el comercio electrónico en base a la conectividad y la velocidad.

Igualmente, se plantearon procesos de modernización en las cadenas de suministro, como el propuesto por Rodríguez (2002), en el cual se realizaba un análisis sobre la planificación de la demanda, con la intención de generar mecanismos de control sobre los objetivos de la organización y la calidad de la información que se disponía. Todo ello, para poder mantener el rumbo, la mejora y la continuidad en el mercado, en un momento dónde Internet comenzaba su irrupción.

Estos conceptos, junto con el crecimiento de Internet (R. A. Lancioni, Smith, & Oliva, 2000) permitieron que los componentes de la cadena de suministro del sector pudiese interactuar más y mejor entre ellos. Precisamente, es lo que comentan Lee & Wang (2004) en el estudio que realizaron sobre el impacto del *e-business*. En este análisis, también añadían que cada vez más clientes u otros agentes podían configurar la información de productos en los sitios web a los que tenían acceso, siendo este fenómeno denominado como *mass customization*. Sin embargo, en este trabajo de Lee & Wang (2004) también se indicaba la necesidad de que los datos, como flujo constante entre los agentes de la cadena, fueran procesados bajo el *Market Intelligence* y en otros casos, también bajo la gestión de la demanda. Es decir, para mejorar en beneficios y posicionamiento de mercado se tenían que aplicar procesos de *Business Intelligence* (BI) y uso de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas, clasificaciones, etc.

2.1 El sector de la distribución

que expliquen el comportamiento de los datos en un determinado contexto o *Data Mining* (DM) para convertir los datos en información. La investigación de Lee & Wang (2004) incluso tuvo en cuenta que estos procesos se combinaron con una nueva dimensión, que surgió en los procesos de negocio denominada monitorización y rendimiento. Es decir, está última dimensión respecto a soporte de post-venta y cumplimiento de pedidos tenía asociadas tareas de seguimiento o *tracking*, pero no era un *tracking* sobre el sitio web de un cliente o de un proveedor.

Yendo más allá, muchos de los estudios citados hablan de la cadena de suministro e incluso de sus agentes o componentes, pero es en el estudio realizado por Murtaza, Gupta & Carroll (2004) donde se realiza un desglose del tipo de relaciones que puede tener un proveedor con otro miembro de la cadena en un mercado virtual, en base a dos dimensiones. Por un lado, la dimensión temporal (a corto o a medio/largo plazo) de la relación, y por otro lado, los tipos de productos o servicios (directos o indirectos) que se proveen a los miembros de la cadena con los que se tiene la relación. De hecho, este tipo de clasificación es importante y puede ser de utilidad, debido a la falta de estudios sobre los proveedores que interactúan en primera instancia con los agentes de distribución.

Asimismo, hay que destacar tres de los trabajos desarrollados por Constantinides (2002, 2004; 2005). El primero de ellos, trata la estrecha relación que el comportamiento de los visitantes virtuales tiene con el marketing. Aunque según el propio Constantinides, los visitantes virtuales consumidores no tienen tanta dependencia de las 4Ps (*Product, Price, Place, Promotion*) del Marketing como de las 4S (*Scope: strategy and objectives, Site: Web Experience, Synergy: integration, System: Technology, Technical Requirements and Web Site Administration*) del marketing web. Continuando con la perspectiva del marketing web, el segundo trabajo de este autor identificaba los factores del comportamiento de los visitantes virtuales, que son consumidores finales, respecto a su experiencia web y también dentro de las categorizaciones que realizaba el estudio ya se detectaba una parte para usabilidad, navegación del sitio, etc. En cambio, en el tercer estudio se utilizaban las categorías generadas para analizar su impacto en el comportamiento de compra, basado en encuestas con un análisis estadístico realizado a posteriori. Cabe destacar que se tuvo

Capítulo 2. Estado del Arte

en cuenta los efectos sobre el vendedor, ya que esto afectaba a la cadena de valor del sector de la distribución, aunque no se especificaba si el vendedor era un mero distribuidor o si también era proveedor. Estos puntos que recogían las investigaciones de Constantinides (2002, 2004; 2005) son un punto de partida fundamental para conocer el comportamiento del visitante virtual en la cadena de suministro del sector y sobre todo en otros participantes de la misma.

Mas adelante y sobre todo a nivel de empresas del sector de distribución en España, González, Mera, Lacoba & Mogollón (2006) crearon un sistema de evaluación en base a un índice de conceptos en el cual se trataba la navegabilidad, contenido, etc. En definitiva, un marco de evaluación por el cual se empezaba a medir la presencia en Internet. Sin embargo, no se indicaba explícitamente el origen del dato, por lo cual era otro caso mas que se basaba en observaciones o algún tipo de encuesta.

Unos años más tarde, se continuó con la parte de experiencias online dentro del comportamiento del visitante virtual. El siguiente paso que se llevó a cabo fue estudiar qué elementos de las experiencias web a la hora comprar, determinaban la elección de un vendedor minorista *online* o no, tal y como recogían Lorenzo, Constantinides, and Gomez-Borja (2009) en su investigación a nivel de estatal. En este caso, se utilizaban encuestas *online* de los consumidores para corroborar las hipótesis que planteaba el propio trabajo a través de análisis estadístico.

No obstante, y a pesar de que es en la parte de la propia distribución (gran consumo y distribuidores) y en los consumidores finales dónde se han realizado la mayoría de las investigaciones, también se han realizado estudios sobre la relación cliente-proveedor. Un ejemplo de ello es el análisis de impacto realizado sobre el comercio electrónico basado en Internet (CEI), en el cual se medían las relaciones cliente-proveedor, teniendo como resultado positivo por un lado, que la distancia entre ambos se reducía gracias a los CEI y por otro lado, lo negativo, que aumentaban los costes de cambio y disminuía el margen de beneficio en los proveedores (Del Águila, Padilla, & Jiménez, 2000).

En este sentido, una mayor aproximación a la tecnológica permitió el desarrollo de

2.1 El sector de la distribución

programas para la integración de los proveedores locales, como el SDP (*Supplier Development Programme*) (Thomas & Barton, 2007) y a través de los mismos se generaba una red local de proveedores. Este tipo de estudios, significaron los primeros contactos con Internet de los proveedores locales, siendo un hito previo a la generación de visitantes virtuales en los sitios web de los proveedores, sobre todo por el empuje de los sistemas de comercio electrónico. De hecho, la mejora en las infraestructuras de red, tecnología, etc. permitió evolucionar el CEI, y tal como afirmó Moyano-Fuentes (2010) "las herramientas basadas en Internet han mejorado significativamente la colaboración e integración de los socios comerciales, permitiendo una vinculación importante entre proveedores y clientes (p.32)".

El establecimiento del foco en la integración de los proveedores en la cadena de suministro y las relaciones entre ellos, se recogen en dos trabajos realizados por Lambert et al (2000) y por Lambert & Schwieterman (2012) respectivamente. El segundo trabajo es una continuación del primero y tiene en cuenta los diferentes niveles de proveedores y el desarrollo de un *framework* o marco de trabajo de métricas, sobre todo financieras y de negocio, que indica que todavía no se tienen en cuenta variables *online* o mejor dicho de comportamiento del visitante virtual.

La relación entre los proveedores de la cadena es fundamental tenerla en consideración, bajo los contextos de los sistemas de comercio electrónico, ya que como se ha comentado previamente su irrupción está asociada a la de Internet.

Por ello, en el contexto del B2B, cuando comenzó la irrupción del comercio electrónico (Gereffi, 2001) aparecieron estudios que analizaban:

- Las virtudes de la cadena y de las propias relaciones de tipo B2B indagando en los tipos de intercambio, reducción de costes de la cadena y unitarios, etc. (Serve, Yen, Wang, & Lin, 2002). No hay que olvidar que los propios sistemas B2B fueron los primeros que tuvieron más influencia en la gestión de la cadena de suministro (Gunasekaran & Ngai, 2004).
- Las compatibilidades entre el propio B2B y otros sistemas como p.e. EDI tanto a nivel tecnológico, como de percepción en las organizaciones. De esta

Capítulo 2. Estado del Arte

manera, se desarrollaban modelos de influencia, ya que no todas las empresas estaban decididas a implantar este tipo de sistemas (Hernández, Jiménez, & Martín, 2006), a pesar de que la conectividad de tipo cliente-proveedor tenía sus ventajas y desventajas (Humphreys, McIvor, & Cadden, 2006).

Esta indecisión situaba al contexto B2B en una coyuntura difícil y era necesario conocer si la compatibilidad de estos sistemas era una cuestión de relaciones o no. En este sentido, la investigación que llevaron a cabo Cullen & Webster (2007) generó un modelo de interacciones de B2B, dónde se podía observar que situación era mas restrictiva para la parte de cliente y cual para la parte proveedora. El modelo ponía hincapié en diversas características como el lugar, la funcionalidad disponible, etc. que se permitían en las relaciones B2B. De hecho, tenía en cuenta si determinados escenarios eran factibles o ventajosos para una masa critica de clientes o proveedores, o si por el contrario estaba más orientado a clientes o proveedores individuales. De esta manera, se facilitaban unas pautas para superar el no atrevimiento de las organizaciones e incluso mejorar el ámbito de las relaciones.

En los contextos de B2C también se han realizado estudios sobre el visitante virtual. Por ejemplo, cabe resaltar el estudio desarrollado por Constantinides, Schepers & Vries (2015), en el cual se realiza un estudio empírico para demostrar el salto existente entre las necesidades y los logros de los visitantes. Aun así este análisis, está muy enfocado a un visitante virtual con el rol de consumidor final y además se basa en cuestionarios y aplicación posterior de estadística para obtener los resultados.

En el B2C, aparte de este tipo de estudios enfocado a redes sociales, también se han desarrollado modelos que con el tiempo han derivado en sistemas que tratan de sugerir, proponer o incluso dar un feedback al visitante virtual. Estos sistemas proporcionan indicadores de valoración, grados de confianza, etc. con el objetivo de persuadir o convencer al visitante virtual para que lleve a cabo la actividad que le interese al sistema.

Sin embargo, en este tipo de sistemas pueden ocurrir recomendaciones negativas en el visitante virtual. De hecho, resulta interesante la investigación llevada a cabo

2.1 El sector de la distribución

por Masterov, Mayer & Tadelis (2015), en la cual se desarrolló un estudio sobre la predicción de experiencias pobres, desagradables, etc. con el objetivo de evitar que los visitantes no dejasen de ser recurrentes. Para ello, se establecía el foco en los mensajes B2S (*Buyer to Seller*), los cuales van desde el visitante virtual con el rol de comprador hasta el de vendedor. A pesar de ello, existen problemas de sesgo en esta clase aplicaciones y suelen implementarse normalmente bajo plataforma B2C (Tadelis, 2016).

A continuación, en la Tabla 2.3 se muestra la aportación bibliográfica con los diferentes estudios de este apartado.

Capítulo 2. Estado del Arte

Tabla 2.3 Retrospectiva de la integración al comportamiento en los canales de venta en la cadena de suministro del sector de la distribución.

Estudio	Autor
Integración, competencia y concentración	Johnson (1999), Christopher and Jüttner (2000), Trappey and Trappey (1998), Lambert and Cooper (2000), Roche et al. (1999)
Impacto del comercio electrónico en las relaciones en base a conectividad y velocidad	Golicic et al. (2002)
Crecimiento de Internet y modernización de la cadena	R. A. Lancioni et al. (2000), Rodríguez (2002)
Impacto del <i>e-business</i> (<i>Business Intelligence y Data Mining</i>)	Lee and Whang (2004)
Desglose dimensional de la relación	Murtaza et al. (2004)
Comportamiento virtual: dependencias con el marketing	Constantinides (2002)
Factores de comportamiento con el consumidor final	Constantinides (2004)
Categorización en el impacto de la compra	Constantinides and Geurts (2005)
Navegabilidad y experiencias	González et al. (2006), Lorenzo et al. (2009)
Comercio electrónico en Internet y tecnología	Del Águila et al. (2000), Moyano-Fuentes (2010)
Integración: niveles de proveedores y desarrollo de variables de negocio	Lambert and Cooper (2000), Lambert and Schwierman (2012)
Contexto B2B	Gereffi (2001), Serve et al. (2002), Gunasekaran and Ngai (2004), Humphreys et al. (2006), Cullen and Webster (2007)
Contexto B2C	Constantinides et al. (2015), Masterov et al. (2015), Tadelis (2016)

Fuente: Elaboración propia

Una vez visto, el recorrido y la retrospectiva de las cadenas de suministro del sector de la distribución, en el siguiente apartado se va poner el foco en la cadena de suministro virtual, dónde se centra esta tesis.

2.2 Cadena de suministro virtual

En este apartado se va a poner el foco en la cadena de suministro virtual del sector de distribución en base a dos ejes. Por un lado, los diferentes escenarios de relación

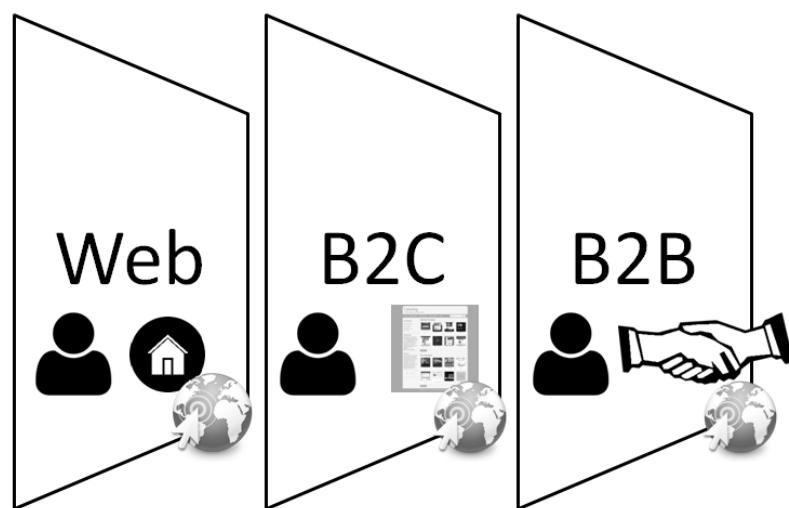
2.2 Cadena de suministro virtual

existentes y por otro lado un conjunto de argumentos sobre las ramas de análisis de comportamiento.

2.2.1 Escenarios de relación

A continuación, se identifican los siguientes escenarios de relación (ver Figura 2.5) en base tanto al esquema de comparativas de las cadenas visto en el apartado anterior (ver Figura 2.4) como a las áreas de conocimiento asociadas a las relaciones en la cadena de suministro, entre proveedores otros componentes de la cadena de suministro y bajo el contexto de los sistemas de comercio electrónico (Castánder et al., 2017):

Figura 2.5: Escenarios de Relación



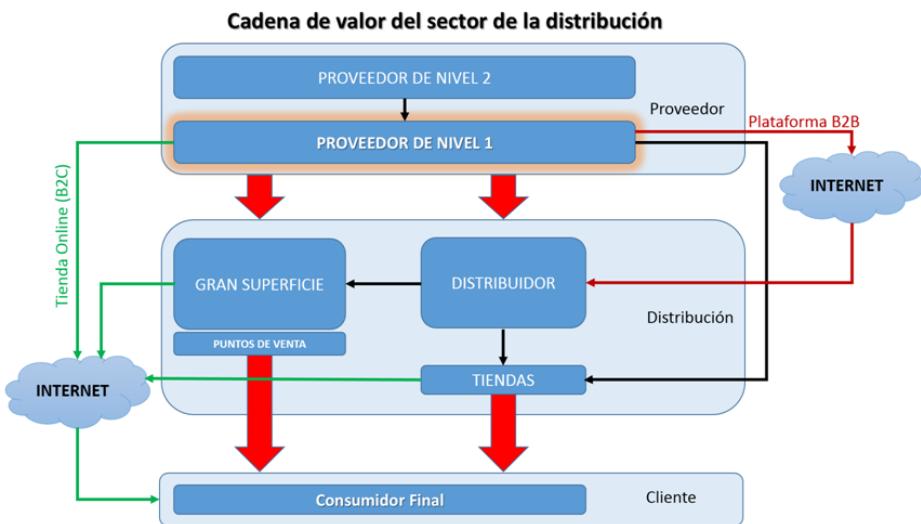
Fuente: Elaboración propia basado en (Castánder et al., 2017)

Tal y como se puede observar en la figura anterior, los sistemas de B2B y B2C, ambos de comercio electrónico, representan la mayoría en los escenarios. De hecho, es el escenario denominado Web, el identificado como el contexto que se utiliza a nivel corporativo y que es necesario para analizar el comportamiento de los visitantes virtuales. Todo debido a las interacciones que se realizan con la parte corporativa de los sitios web. En este sentido, para situar mejor dichos escenarios se describe el siguiente esquema de la cadena de valor del sector de la distribución,

Capítulo 2. Estado del Arte

en el cual se muestra como los escenarios de B2B y B2C tienen un papel significativo, tanto para los distribuidores como para los proveedores y como el escenario web de acceso a la web corporativa se puede realizar a través de Internet:

Figura 2.6: Cadena de Valor del sector de la distribución



Fuente: Elaboración propia

2.2.2 Consideraciones sobre las ramas de análisis de comportamiento

Una vez visto los escenarios de relación y como se ubican en la cadena de valor del sector de distribución, con el propósito de organizar y concretar los análisis a realizar en los diferentes escenarios de relación es necesario definir el conjunto de las ramas de análisis de comportamiento dentro del proyecto de investigación.

Entendiendo como **rama de análisis de comportamiento del visitante virtual** a un ámbito del caso de estudio, que abarca un tema determinado asociado a los escenarios de relación.

Aun así, teniendo en cuenta el trabajo realizado por Castánder et al (2017) sobre la metodología para el comportamiento de los visitantes virtuales en los proveedores del sector de la distribución y las observaciones que se vayan a realizar tras conversar con los diferentes agentes de los casos de estudio, hay que tener en cuenta

2.3 Conocimiento del comportamiento del visitante virtual a través del *Web Analytics*

las siguientes consideraciones a la hora de definir las ramas de análisis de comportamiento:

- Identificación del agente o agentes de la cadena de valor del sector de la distribución desea el **proveedor de primer nivel estrechar y mejorar su relación** para poder disponer de un mayor control de los canales digitales.
- Las **relaciones de los proveedores** con otros agentes de la cadena de valor del sector de la distribución en base a las **dimensiones de tiempo y de producto/servicio**. En lo referente a la dimensión de tiempo para no pasar por alto el control del canal digital, el acceso al producto, etc. y en la dimensión de producto/servicio para no pasar por alto la evolución IT del producto, la captación y vigilancia, etc.

Una vez visto los escenarios de relación y las consideraciones sobre las ramas de análisis de comportamiento dentro de este apartado de la cadena de suministro virtual, a continuación se describe el apartado sobre el *Web Analytics* necesario para obtener el conocimiento sobre comportamiento de los visitantes virtuales.

2.3 Conocimiento del comportamiento del visitante virtual a través del *Web Analytics*

En el apartado anterior de la cadena de suministro virtual se definían los escenarios de relación y un conjunto de consideraciones para que se pudieran definir ramas de análisis de comportamiento. Sin embargo, dicho comportamiento concretamente del visitante virtual se tiene que obtener de alguna forma y para ello está el *Web Analytics*. En este sentido, según la *Digital Analytics Association* (DAA) (antes *Web Analytics Association* (Burby, Brown, Committee, et al., 2007)) (Association, 2008) se entiende por *Web Analytics* cómo "la medición, recopilación, análisis e informe de datos de Internet para el propósito de entender y optimizar el uso de la web (p.3)". Un seguimiento por el cual cada vez más empresas están interesadas (Jarvinen & Karjaluoto, 2015). Por consiguiente, en este apartado se definen las tres fases que van a permitir disponer los datos y convertirlos en información para comprender del comportamiento del visitante virtual.

Capítulo 2. Estado del Arte

2.3.1 Obtención y acceso a las fuentes de datos del comportamiento

El comportamiento del visitante virtual ha necesitado el soporte de un proceso TIC denominado (KDD) *Knowledge Discovery in Databases*, que permite el descubrimiento del conocimiento y la información útil dentro de los datos que están contenidos en los repositorios de información (Fayyad, Piatetsky-Shapiro, & Smyth, 1996). Dentro de estos procesos, la etapa de recogida de datos y el estudio de tipos de variables que tienen que contener tal información juegan un papel fundamental. Es decir, es importante saber la evolución de los sistemas que han obtenido la información desde la navegación del visitante web. Para cualquier proveedor de información online y analista web existen en esencia tres puntos que pueden tomar el pulso del consumidor digital: *logs* de acceso o entrada, explicación interna y el *tracking & tagging*. Los *logs* de acceso fueron durante años el punto de referencia para usar estos sistemas de recogida de datos. Tales sistemas eran capaces de registrar explícitamente el comportamiento de navegación de los visitantes de los sitios web (Srivastava, Cooley, Deshpande, & Tan, 2000; Naresh Kumar Kar, 2015). Pero mucho ha cambiado desde la última década. Las páginas web ya no son simples archivos de texto con unas cuantas fotos auxiliares y *scripts*. La página web moderna está mejor pensada como una confusión compleja e híbrida del código del programa y del contenido que se procesa por la máquina virtual del navegador web. El resultado es que el *log* de acceso HTTP ya no es una idea rentable y fiable en la corriente de la conciencia del usuario en línea (Clark, Nicholas, & Jamali, 2014).

Por todas estas razones, hoy en día en lo relativo a los **sistemas de recogida de información**, los sistemas de *web tracking* como *Google Analytics*, *Piwik*, etc., están reemplazando los sistemas de *logs* de acceso web y su impacto podría ser más fuerte si estos *trackers* consiguen la sustitución de las encuestas en línea en la fase de recuperación de datos el visitante virtual. Estos sistemas de *tracking* permiten registrar, mediante la inserción de un *script* en el sitio web, cierta información que se almacena en un conjunto de variables de diferentes tipos, cuando las visitas se producen en el sitio web. De hecho, entre dichos sistemas de *tracking* por un lado está *Google Analytics*, el cual está muy ligado a una perspectiva de marketing

2.3 Conocimiento del comportamiento del visitante virtual a través del *Web Analytics*

de comportamiento *online*. Por otro lado, *Piwik*¹, que destaca por ser *open-source* a pesar de no tener una aprobación tan unánime como *Google Analytics* (Clark et al., 2014).

Tabla 2.4 Componentes de comercio electrónico para los sistemas de recogida de información

Sistema de recogida de información	Componente de comercio electrónico
<i>Kissmetrics</i>	<i>Kissmetrics ECommerce</i> *
<i>Clicky</i>	<i>Clicky ECommerce Plugin (Magento, PrestaShop)</i> **
<i>Google Analytics</i>	<i>Google Analytics ECommerce</i> ***
<i>Piwik</i>	<i>Magento, PrestaShop, ZenCart, WooEcommerce, etc.</i> ****

* <https://www.kissmetrics.com/ecommerce/>

** <https://clicky.com/help/apps-plugins#magento>

*** <https://developers.google.com/analytics/devguides/collection/analyticsjs/ecommerce>

**** <http://piwik.org/integrate/>

Fuente: Elaboración propia

Como se puede observar en la Tabla 2.4 tras realizar una revisión del mercado, se ha hecho hincapié en los sistemas de *web tracking* principales en la actualidad. Entre ellos están *Google Analytics* y *Piwik*. La diferencia es que el primero ofrece un único componente oficial para comercio electrónico pudiendo disponer del componente estándar y del mejorado. En cambio, en el caso de *Piwik* se dispone desde la página oficial toda una gama de componentes para comercio electrónico, siendo *Magento* uno de los principales. Además, *Piwik* ofrece flexibilidad respecto al alojamiento, ya que aparte de la opción *cloud* permite disponer la información en servidores propios.

En este sentido, y con el objetivo de ahondar en los pros y contras de cada sistema de *web tracking*, se va mostrar una comparativa con las ventajas y desventajas de

¹piwik.org

Capítulo 2. Estado del Arte

cada herramienta y de otro grupo de herramientas de *tracking* de menor entidad o que están más orientadas a entornos de producción (estas solamente tienen la parte de analítica, o no tienen la parte de comercio electrónico y por lo tanto se han encapsulado en el grupo de otros).

2.3 Conocimiento del comportamiento del visitante virtual a través del Web Analytics

Tabla 2.5 Comparativa de las herramientas de *Web Tracking* a utilizar para el origen de *Web Analytics*.

Web Tracking	Ventajas	Desventajas
Clicky	Dispone de un API para acceder a los datos aparte de los informes o paneles. Realiza las analíticas en tiempo real y su interfaz de informes es clara y sencilla.	No tiene un conjunto de variables propias de comercio electrónico aunque tiene <i>plugins</i> para tiendas como Magento, etc. La versión <i>free</i> está limitada a un sitio web y a un total de páginas vista por día. Los datos están en su servidor.
Google Analytics	Es la más extendida y utilizada a nivel mundial. La opción de comercio electrónico y comercio electrónico mejorado dispone de gran cantidad de variables.	La versión <i>free</i> está limitada. Los datos están en su servidor.
Universal Analytics	Dispone de un API para acceder a los datos aparte de los informes o paneles.	Su estructura de <i>tracking</i> de <i>user</i> , <i>session</i> y <i>hit</i> es diferente a otros sistemas.
KissMetrics	Dispone de una ampliación llamada <i>Universal Analytics</i> que complementa la versión tradicional con multiplataforma, variables personalizadas, etc.	
	Dispone de información sobre CLV, <i>cohorts</i> , campañas, etc.	
KissMetrics	Tracking bastante completo con muchas opciones de panel de informes e incluso opciones de segmentación.	Es de pago y de precio bastante elevado.
	Dispone también de opción de comercio electrónico.	Los datos están en su servidor. Las métricas de tiempo real no son muy precisas.
Open Web Analytics	Es totalmente <i>free</i> .	La instalación no está sencilla es más manual que otros sistemas.
	Permite trackear múltiples sitios web.	En el sitio web de este sistema no está disponible el esquema de las variables. Accediendo al <i>back-end</i> del mismo si se pueden obtener a las partes que lo componen.
Piwik	También dispone de acceso a los datos mediante API. Dispone de <i>tracking</i> de comercio electrónico.	El mantenimiento respecto a gestión de <i>bugs</i> depende de la comunidad y del autor que publica las versiones bajo licencia GPL (<i>General Public License</i>).
	Muchas de las características de Google Analytics están implementadas en este sistema.	En su versión <i>free</i> no dispone de un apartado de comercio electrónico tan potente como Google Analytics.
Piwik	Es <i>free</i> si el <i>host</i> está en nuestro propio servidor.	Su panel de informes no está tan organizado como el de otros sistemas.
	Permite trackear múltiples sitios web. Dispone de <i>tracking</i> de comercio electrónico y de acceso a los datos mediante API.	El mantenimiento respecto a gestión de <i>bugs</i> depende del propio <i>Piwik</i> que va actualizando el sistema. Aun así existen versiones LTS.
Otros (<i>Adobe Analytics</i> , etc.)	A nivel de panel de informes este grupo de herramientas son bastante completas.	La mayoría de este grupo o están muy orientadas a entornos de producción por lo cual tienen un precio considerable de licencia, o son totalmente <i>free</i> pero no tienen un mínimo de funcionalidades y siempre suelen no tener alguna como por ejemplo: <i>tracking</i> de comercio electrónico.

Fuente: Elaboración propia basada en (Fernando, 2014; Ciotti, 2015)

Capítulo 2. Estado del Arte

Una vez, que han sido presentados los sistemas de huella digital, a continuación, se va a exponer un estudio sobre los indicadores o variables que son recogidos por los sistemas de *tracking*. De hecho, existen muchos tipos de datos que pueden ser utilizados en el *Web Mining* (ver punto 2.3.3). Tales datos están clasificados en los siguientes tipos (Srivastava et al., 2000):

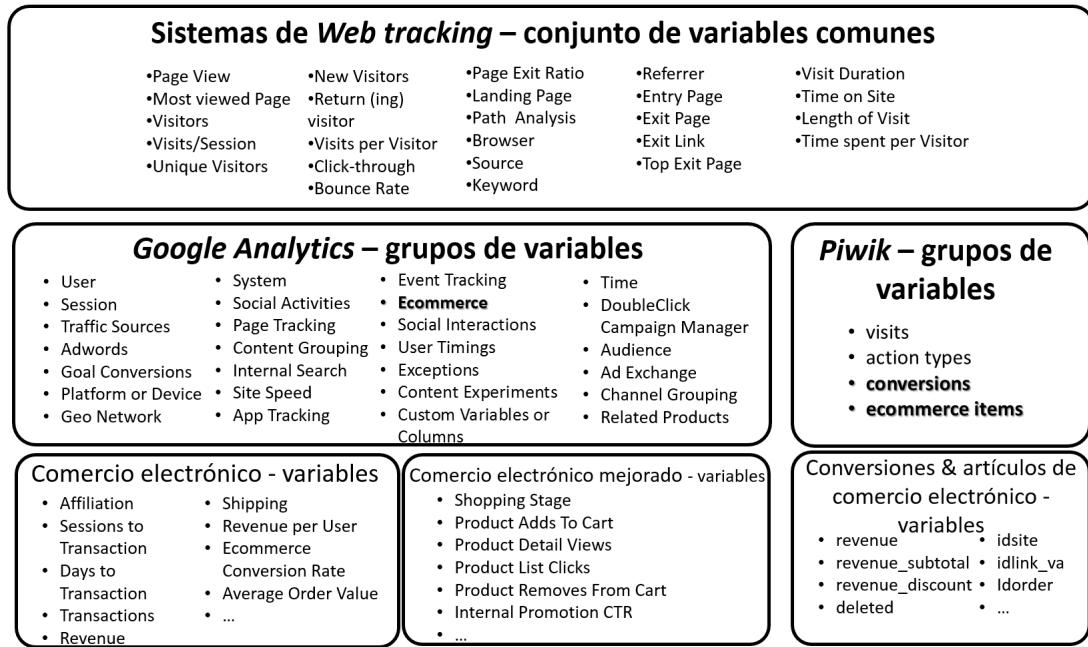
1. Contenido: los datos reales de las páginas web (texto, imágenes, . . .).
2. Estructura: los datos que describen la organización del contenido.
3. Patrones de uso de las páginas web (direcciones IP, referencias de página, fecha y tiempo de los procesos).
4. Perfil del usuario: datos que proporcionan información demográfica sobre los usuarios del sitio web.

No obstante, y siendo un buen punto de partida el llevar a cabo el estudio de las variables, la clasificación mencionada anteriormente hoy en día podría ser más flexible. Especialmente, en el caso del comportamiento del visitante virtual, donde el origen de datos ha sido típicamente el uso, existe el cambio de tener influencia en otras categorías tales como el perfil del usuario y en una capa de toma de decisiones basada en los datos sobre el contenido y la estructura respectivamente.

Por ello, a la hora de llevar a cabo cualquier proyecto que se base en *Web Analytics* es necesario saber las herramientas de recogida de datos y también conocer las variables que se pueden disponer (Hausmann, Williams, & Schubert, 2012; Järvinen & Karjaluoto, 2015; Vellingiri, Kaliraj, Satheeshkumar, & Parthiban, 2015). A continuación, se muestra un **estudio sobre las variables que pueden obtener los sistemas de recogida de información**. Este análisis se divide en tres bloques principales y un esquema de relaciones entre métricas web y métricas de *e-business*.

2.3 Conocimiento del comportamiento del visitante virtual a través del Web Analytics

Figura 2.7: Estudio de variables - Recorrido desde las genéricas hasta las de comercio electrónico de *Google Analytics* y *Piwik*



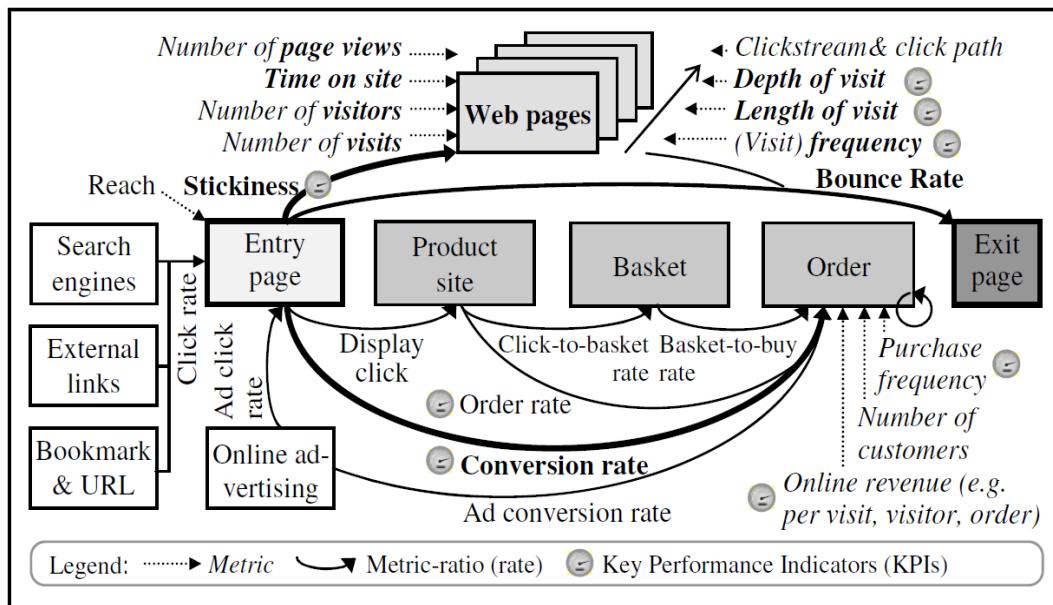
Fuente: elaboración propia basada en (Hausmann et al., 2012; Jarvinen & Karjaluoto, 2015; Vellingiri et al., 2015; Google, 2016; Piwik, 2016)

Tal y como se puede observar en la figura anterior, el primer bloque contiene aquellas variables comunes para cualquier sistema de *web tracking*, el segundo la diferencia entre las variables disponibles de dos de los sistemas más importantes como son *Google Analytics* y *Piwik* y el tercero las variables de ambos, que están orientadas a comercio electrónico.

Una vez visto el recorrido de las variables, el estudio de variables continúa con el esquema de las relaciones entre las métricas *e-business* y las métricas web, con el objetivo de poder observar la ubicación de las variables en visitas virtuales a comercios electrónicos.

Capítulo 2. Estado del Arte

Figura 2.8: Estudio de variables - Relaciones entre métricas web y de e-business



Fuente: (Fasel & Zumstein, 2009)

Como se puede observar en la Figura 2.7 y Figura 2.8, el estudio de variables se ha centrado en aquellas relacionadas con el comercio electrónico, teniendo en cuenta que tienen que ser unas variables que tienen que estar enfocadas con las necesidades del sector de la distribución (ver punto 2.1.2). Las variables que aparecen en este estudio tienen la opción de ser adquiridas a través de APIs. Aun así, y sobre todo en el caso de que no se pueda llegar a las variables requeridas o a la granularidad deseada, para la integración de dichas variables es recomendable utilizar herramientas como las que se describen en el siguiente capítulo.

2.3.2 Integración de los datos de comportamiento

Para llevar a cabo la integración de los datos, es necesario establecer un conjunto de pasos que permitan el establecimiento de los datos, en una estructura común a pesar del origen de los mismos. En este sentido, el BI resulta fundamental para llevar a cabo tal cometido. Un BI que a menudo es referido como las técnicas, tecnologías, sistemas, prácticas, metodologías y aplicaciones que analizan los datos críticos del negocio para ayudar a la organización a comprender mejor sus actividades, mer-

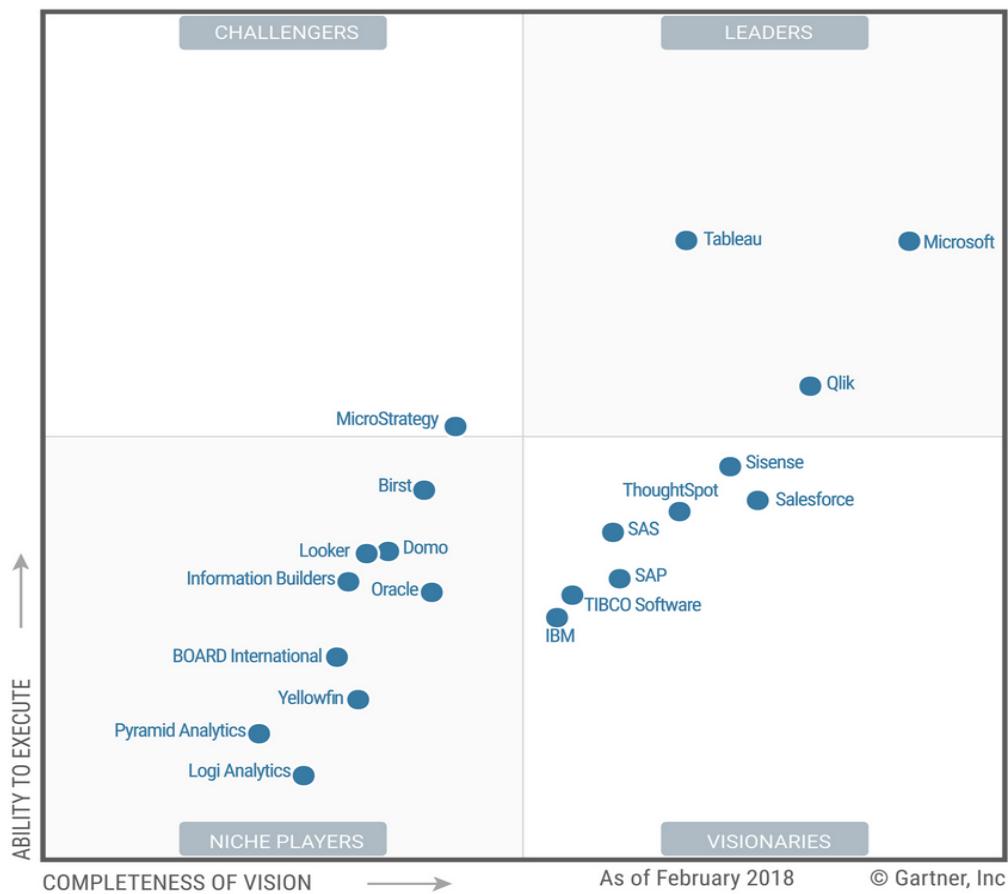
2.3 Conocimiento del comportamiento del visitante virtual a través del Web Analytics

cados y tomar decisiones de negocio a tiempo. (Watson & Wixom, 2007; Negash & Gray, 2008; Chaudhuri, Dayal, & Narasayya, 2011; H. Chen, Chiang, & Storey, 2012). Por lo cual, es muy importante escoger la herramienta que lleve a cabo el BI.

En el mercado existe una amplia gama de organizaciones dedicadas a BI con un catálogo amplio y diverso de *software* o sistemas de información que soportan los procesos de herramientas de BI. Una muestra se puede ver en el informe publicado por *Gartner* (2015) consultado regularmente como "Cuadrante Mágico" de *Gartner*. En este documento, se contempla un "Cuadrante Mágico" que facilita la colocación gráfica de cuatro tipos de proveedores de tecnología en los mercados: los líderes, los visionarios, empresas orientados a nichos y por último los rivales o retadores. Estos se clasifican sobre la base de su crecimiento y su capacidad para diferenciar sus productos.

Capítulo 2. Estado del Arte

Figura 2.9: Cuadrante Mágico de *Gartner* para BI & Analytics Platforms 2018



Fuente: (Aziza, 2018)

De la figura anterior destaca *Birst* y *MicroStrategy* a nivel de centralidad de cara conseguir un mejor equilibrio en comparación con el resto. También hay que remarcar la posición de liderazgo que ocupan *Microsoft*, *Qlik* y *Tableau*. No obstante, se tiene que tener en cuenta que estos análisis, como se muestra arriba, no siempre pueden satisfacer las necesidades de las organizaciones en materia de elegir definitivamente una herramienta de BI.

Por lo tanto, es necesario complementarlo con otro más centrado en las PYMES (Pequeñas y Medianas Empresas) y en el cual, se debe llevar a cabo una evaluación comparativa entre las herramientas de BI que ofrecen las empresas. Siempre desde dos perspectivas: el costo y las especificaciones técnicas que pueden ofrecer

2.3 Conocimiento del comportamiento del visitante virtual a través del Web Analytics

(Castánder, Gerrikagoitia, & Alzua-Sorzabal, 2013). Sin embargo, ambas perspectivas, tarde o temprano se verán influidas por la evolución que las herramientas de BI están teniendo, hacia soluciones *Big Data*. Las herramientas de BI en la etapa de integración tiene como resultado el **almacén de datos**. Este tipo de BD es el componente principal de los procesos de BI, es su DW. Esta estructura permite organizar la información almacenada en tablas de hechos y en dimensiones, proporcionando una clara visión entre lo que ha ocurrido en un momento o tiempo determinado y lo que meramente está definido, es decir, sin que esté asociado a un tiempo concreto. Estableciendo el mecanismo que permite responder a las preguntas que se realizarán en el diseño dimensional. Retomando la información que suelen disponer los DW, a priori suele ser analítica ya que los DW suelen tener como fuente de información BBDD o fuentes operacionales. En el caso de este proyecto de investigación el DW de comportamiento contendrá los datos proporcionados por los sistemas de recogida de información. Hay que recordar que la estructuración de los datos provenientes de los orígenes de los mismos, a través de los servicios de integración permitirá que los datos puedan convertirse en información debido al tratamiento, limpieza, etc. que se pueda dar.

2.3.3 Análisis de los datos de comportamiento

En este punto se describe la etapa más importante de las que componen el comportamiento del visitante virtual. El análisis o estudio de los datos es fundamental para comprender en este caso al visitante virtual. A continuación, se realiza un recorrido por las diferentes áreas dentro del análisis de los datos: estadística, *Data Mining* y *Web Mining*. Comenzando por los estudios estadísticos hasta las técnicas de *Data Mining* dentro del *Web Mining* pasando por el propio *Data Mining*.

2.3.3.1 Método Estadístico

En un principio los primeros análisis fueron meramente estadísticos. Es decir, sobre los datos se aplicaban cálculos matemáticos las medias, el rango intercuartil (IR), etc. con el objetivo de conocer la evolución de los valores que podía tener una variable determinada. Yendo más allá, se empezaron a realizar comparaciones entre los valores de las variables y los campos calculados como por ejemplo la media. De

Capítulo 2. Estado del Arte

esta manera, aparecieron las varianzas, covarianzas, desviación estándar, etc. con las cuales se quería saber más sobre la evolución de los valores del dato.

De cualquier forma, estos tipos de análisis solamente describen a las variables. De hecho, a medida que ha ido evolucionando la organización de los datos, con la aparición de estructuras, clases, etc. se han ido requiriendo otro tipo de análisis para no poner el foco en lo descriptivo, sino en predecir los valores futuros y/o en realizar clasificaciones o agrupaciones de los mismos.

Aun así, hoy en día es necesario como primer paso o paso previo a estudiar los datos, el realizar un análisis descriptivo de los datos. ¿Por qué? el motivo es simple, es necesario conocer cada variable antes de cometer un análisis. Es decir, conocer su rango intercuartil, saber si son datos normalizados y conocer sus *outliers* o datos sobre salen de su rango. Estos dos aspectos, la normalización del dato y la detección de *outliers* es fundamental y repercute en los análisis a realizar a posteriori. Por ello, el análisis de los datos dió un paso más y se comenzó a gestar el DM.

2.3.3.2 Data Mining

Dentro de los procesos KDD, la etapa de aplicar métodos de DM (*Data Mining*) juega un papel importante (Fayyad et al., 1996). De hecho, en los últimos 10 años el campo del DM pasó por una gran transición. El cual se ha recuperado desde el pico exagerado de la era "punto com" y ahora se está convirtiendo en parte del proceso de negocio principal (Piatetsky-Shapiro, 2007). En este sentido, para llevar a cabo el DM existen otros procesos como por ejemplo SEMMA y CRISP-DM. Este último, tiene similitudes con KDD y se resume en el siguiente conjunto de pasos definidos en principio por Chapman (Azevedo, 2008):

1. *Business understanding*
2. *Data understanding*
3. *Data preparation*
4. *Modeling*

2.3 Conocimiento del comportamiento del visitante virtual a través del *Web Analytics*

5. *Evaluation*

6. *Deployment*

Según Azevedo (2008) "CRISP-DM hace de guía para las personas que quieren saber como puede ser aplicado en la práctica el DM en sistemas reales (p.5)". Aun así, aparte del proceso el DM ha ido evolucionando según los tipos u origen de los datos (locales, web, imágenes, textos,...).

Por ello, han surgido nuevas áreas de investigación como el análisis de la red social (enlaces), *Multimedia Mining* y *Web Mining* (Piatetsky-Shapiro, 2007). En este sentido, el *Web Mining* es el punto en el que se va a centrar esta parte del estado del arte.

2.3.3.3 **Web Mining**

Es el concepto que reúne todas las técnicas, métodos y algoritmos utilizados para extraer información y conocimiento a partir de datos procedentes de los sitios web. En otras palabras, los esfuerzos DM asociados a la Web se consideran *Web Mining* y se pueden dividir en tres clases: *content mining*, *usage mining* y *structure mining* (Srivastava et al., 2000; Kaur & Khanna, 2014; Sakthipriya, Srinaganya, & Sathia-seelan, 2015).

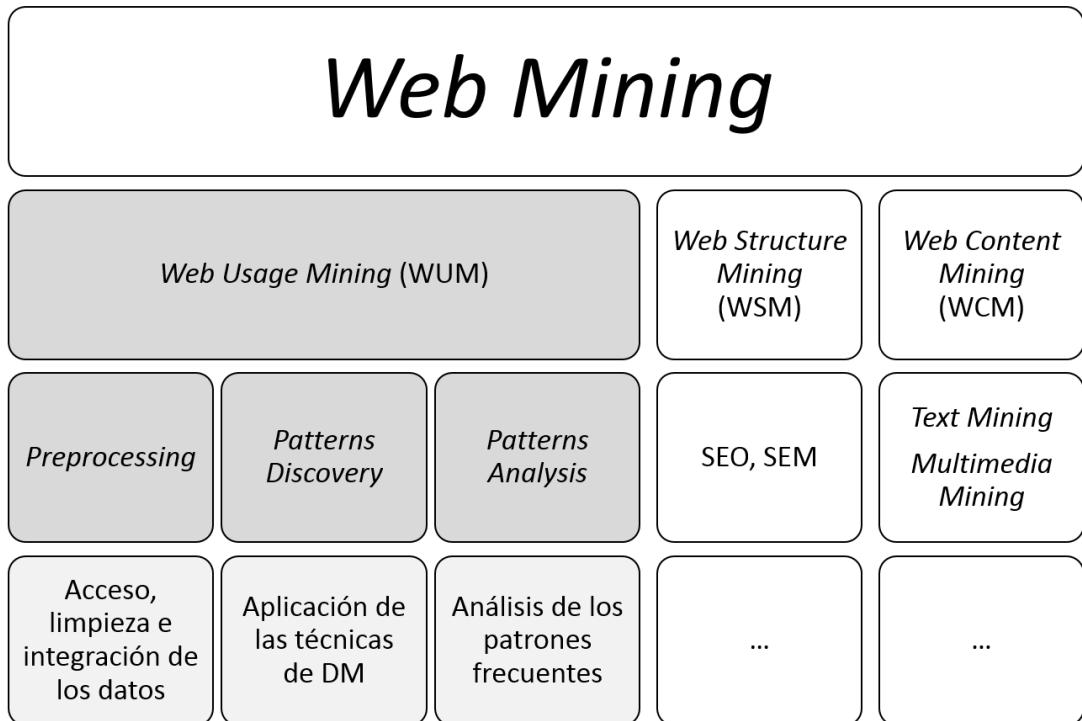
Algunos de ellos apuntan a analizar el comportamiento del usuario con el fin de mejorar continuamente la estructura y el contenido de los sitios que se visitan. El *Web Mining* en sí se divide en tres categorías principales: contenido; la estructura y el uso. Dependiendo del tipo de datos del sitio web para procesar (Velásquez & Donoso, 2010). En este sentido, la utilización del *Web Mining* es la aplicación de técnicas de DM para descubrir patrones de uso de datos de la Web, a fin de entender y servir mejor a las necesidades de las aplicaciones basadas en la Web. WUM (*Web Usage Mining*) consta de tres fases, a saber, pre-procesamiento, descubrimiento de patrones y análisis de patrones (Srivastava et al., 2000).

A continuación, se muestra de manera gráfica el esquema que sigue el *Web Mining*. Es decir, las áreas que lo componen y comentadas anteriormente. Y dentro

Capítulo 2. Estado del Arte

del WUM que etapas dispone y estas como llegan hasta el DM y otras tareas en relación a los datos.

Figura 2.10: Esquema del *Web Mining*



Fuente: Elaboración propia basada en (Agarwal, Khan, & Dhall, 2010; Aggarwal, 2014; Sakthipriya et al., 2015; Sulbhewar & Raut, 2015)

Tal y como se muestra en la Figura anterior, el WUM está compuesto por las fases de pre-procesamiento de los datos, descubrimiento de patrones y análisis de los mismos. Cada una de ellas supone llevar a cabo un conjunto de tareas o actividades que se comentan a continuación:

1. **Acceso, limpieza e integración de los datos**, son las actividades que están bajo el pre-procesamiento de los datos. Es decir, disponer de un acceso estable a los datos, aplicarles técnicas estadísticas descriptivas para la limpieza de los mismos y llevar a cabo una integración de los mismos mediante las soluciones tecnológicas de BI y/o de *Big Data*, que culminen en un almacén de datos disponible para el análisis.

2.3 Conocimiento del comportamiento del visitante virtual a través del *Web Analytics*

2. **Aplicación de las técnicas de DM**, es la actividad sujeta al descubrimiento de patrones. Para ello, es necesario en primer lugar decidir que se quiere realizar: predecir, clasificar, agrupar,... A continuación, decidir la técnica o método de DM más apropiado o si se tiene que aplicar una combinación de las mismas y finalmente aplicarlas sobre los datos.
3. **Ánalysis de los patrones frecuentes**, es la tarea asociada al análisis de patrones a través de la cual se quiere identificar las pautas habituales o usuales en los resultados de la aplicación de las técnicas de DM. Para ello, aparte de las técnicas de visualización gráfica que representan la forma más heurística de realizar el análisis, es necesario utilizar *Frequent Pattern Mining* (Aggarwal, 2014).

En relación con el CRM, el concepto de *Web Mining* se mantiene y va más allá del descubrimiento de patrones. De hecho se utiliza para ayudar a cuantificar el éxito de una campaña de marketing. Y más a nivel general para evaluar la efectividad de un sitio web particular y comprender el comportamiento del cliente (Agarwal et al., 2010). Un comportamiento que en lo referente a la aplicación que tiene el *Web Mining* sobre el CRM se traduce en mejorar el tiempo que permanecen los clientes en el sitio web; la retención de los clientes y la adquisición de los nuevos (Q. Liu & Zhang, 2013).

2.3.4 Técnicas de *Data Mining* disponibles en el *Web Mining*

Con el objetivo de profundizar en el conocimiento de las técnicas que pueden llevar a cabo el análisis de los datos de comportamiento, se comentan brevemente algunos de los métodos de DM utilizados en *Web Mining*:

2.3.4.1 ANN (*Artificial Neural Network*) & SOM (*Self-Organizing Mapping*)

La técnica ANN es una aproximación de funciones universales y se han utilizado con éxito en muchos campos de investigación, como el control del tráfico aéreo, reconocimiento de voz y los personajes, el diagnóstico y la investigación médica, la predicción del tiempo, etc.

Capítulo 2. Estado del Arte

Este método representa una arquitectura de neuronas que reciben las variables de entrada y son responsables para el procesamiento de ellas para proporcionar descendencia por una o más variables. Una implementación de este método es SOM (Lagus, Kaski, & Kohonen, 2004; Bloom, 2004; Wehrens & Buydens, 2007) cuyas principales características son la encapsulación de la funcionalidad y la visualización enriquecida.

2.3.4.2 Clustering

Esta técnica se sitúa dentro del *Pattern Discovery* en el WUM (Rathod, Prajapati, & Khanna, 2015; Sakthipriya et al., 2015). Es una de las más comunes tanto del *Data Mining* como del *Web Mining*. Dentro esta técnica uno de los métodos más destacados es el *K-Means*.

Este se basa en la cálculo de los grupos o *clusters* dentro de un conjunto de datos. Es decir, en primer lugar se estable como medida el número esperado de los grupos (k). A continuación, cada uno de los grupos estará representado por un elemento cuyas características se parecen más a su conjunto (obtenido por una medida de similitud), denominado centroide. El método realizará diferentes iteraciones y en cada una de ellas los centroides se irán recalculando hasta que el método finalice, siendo ese caso cuando se establecen los centroides definitivos de los grupos. Además, debido al alto rendimiento que propicia este método es posible repetir varias veces con diferentes parámetros (Velásquez & Donoso, 2010).

2.3.4.3 LDA (*Linear Discriminant Analysis*) & LR (*Logistic Regression*)

Ambos métodos, LDA (Fisher, 1936; Trevino & Daniels, 1995; Kim, Kim, Kim, Ye, & Lee, 2000) y LR (Suh, Noh, & Suh, 1999; Kay Olsen, Warde, & Martens, 2000; Hastie, Tibshirani, & Friedman, 2009) se utilizan para resolver problemas de clasificación en el ámbito de las estadísticas. LDA aplica un método LDF (*Linear Discriminant Function*) que toma ventaja de centroides de una manera simple, discrimina entre dos grupos. Este hecho no significa que sea una técnica más eficiente y preciso que LR (Harrell & Lee, 1985; Hastie et al., 2009); ya que esta utiliza la

2.3 Conocimiento del comportamiento del visitante virtual a través del *Web Analytics*

probabilidad de un resultado dicotómico en relación con un conjunto de variables potencialmente independientes.

2.3.4.4 **CART (Classification And Regression Tree) & RF (Random Forest)**

Ambas técnicas, basan su arquitectura en un esqueleto con un formato de árbol cuyo propósito es realizar una clasificación o regresión de los datos (Breiman, Friedman, Olshen, & Stone, 1984). CART (Hastie et al., 2009) implica la aplicación de tres pasos:

1. Se construye un extenso árbol denominado "*maximal tree*" que describe con cuidado el conjunto de entrenamiento y que crece a través de un proceso de separación binaria.
2. El árbol se poda y deriva en una serie de árboles de menor complejidad.
3. Se selecciona a través del método de CV (*Cross-Validation*) el árbol cuyo tamaño es óptimo.

La mayor aplicabilidad de CART ha estado en el campo de la meteorología, como pronosticadores del tiempo (Davis, Elder, Howlett, & Bouzaglou, 1999), en las ciencias sociales (Fu, 2004) y en la medicina (Kurt, Ture, & Kurum, 2008). A diferencia del método anterior, en RF está disponible un conjunto de árboles para clasificar. El proceso mejora si cada árbol es diferente y si esta aleatoriedad entre los árboles es proporcionada por diversas muestras independientes de los datos de entrenamiento y considerando un subconjunto de los atributos de los datos seleccionados al azar (Breiman, 2001).

2.3.4.5 **AR (Association Rules)**

El propósito de este método es generar un gran número de normas y extraer las que son útiles objetiva y subjetivamente. Objetivamente, a través de un análisis estadístico de los datos.

Sin embargo, a pesar de que es un procedimiento que normalmente se utiliza en el análisis de "*market-basket*", también se prefiere eliminar el conocimiento de los

Capítulo 2. Estado del Arte

patrones de comportamiento de usuarios (Sánchez, Vila, Cerda, & Serrano, 2009); ya que el algoritmo "*Min you*" funciona bien en situaciones de bajo apoyo y alta confianza(Lai & Cerpa, 2001).

2.3.4.6 SVM (Support Vector Machines)

Es una técnica poderosa que permite la clasificación y la regresión sobre los datos (Vapnik, 1998). Para ello, es responsable de la generación de un hiperplano que permite la clasificación a un lado de los datos que cumplen con la condición y al otro lado con aquellos que no la cumplen. La separación propuesta de los datos es llevada cabo obteniendo la mayor distancia entre los datos que cumplen la condición y los más cercanos que no la cumplen (R.-C. Chen, Chiu, Huang, & Chen, 2004).

Su sencillez, en el momento preciso de las clasificaciones lineales, y su capacidad para jugar con una amplia gama de variables, hacen que este sea un método atractivo para comprender el comportamiento del visitante virtual.

2.3.4.7 GA (*Genetic Algorithms*)

Inspirado en la evolución natural, GA pretende proporcionar una mejor solución a medida que transcurre el tiempo; ya que la probabilidad de supervivencia de los miembros más fuertes de la población es superior a la de los miembros más débiles. Normalmente, las nuevas generaciones se producen por el cruce de dos miembros; pero también, pueden ocurrir mutaciones en los miembros que faciliten la diversidad de la población (Duman & Ozcelik, 2011). Cabe señalar que se trata de un método que soporta fácilmente la complementariedad en el algoritmo evolutivo; permitiendo este hecho, mejorar la velocidad del proceso.

Yendo más allá, hay que tener en cuenta que fueron realizados enormes progresos en refinar los algoritmos AR (aunque las buenas aplicaciones son pocas todavía), y los nuevos algoritmos importantes tales como SVM han sido ampliamente aceptados. Además, el *software* de DM de *open-source* como Weka ¹ y R (lenguaje de

¹<http://www.cs.waikato.ac.nz/ml/weka/>

2.3 Conocimiento del comportamiento del visitante virtual a través del *Web Analytics*

programación estadístico y de análisis) llevaron a un mayor uso de las herramientas de DM (Piatetsky-Shapiro, 2007).

Una vez visto el apartado anterior, a continuación se van a enumerar un conjunto de herramientas que se van a poder utilizar para llevar el análisis de datos a realizar en el proyecto de investigación:

1. Suite de BI: *Microstrategy*
2. Lenguaje de programación de análisis y estadística: R
3. *Pack* de Integración de R con la suite de BI: *Microstrategy*
4. *Packs* de repositorios de R (p.e. *CRAN*), para el análisis y evaluación de datos

Sigue un orden y trabaja en secuencia.

Anónimo

CAPÍTULO

3

Metodología de la investigación

En este capítulo se describe la metodología a seguir en el proyecto de investigación de cara a cumplir el objetivo general del proyecto que es: **desarrollar un modelo de comportamiento del visitante virtual, en el marco de los proveedores de primer nivel en el sector de la distribución, para el caso de estudio de ORBEA S.Coop..**

En el primer apartado se van a analizar los diferentes enfoques de investigación existentes, como base para justificar el tipo de enfoque metodológico, el tipo de estudio (descriptivos, exploratorios,...), la estrategia (experimento, estudio de casos,...) y el tipo de estudio de casos (caso único, múltiples). Concluyendo que el que mejor se adapta al propósito de esta investigación, es el enfoque cualitativo con tipo de estudio exploratorio, aplicando una estrategia de estudio de casos y concretamente de caso único.

En el segundo apartado se plantea el diseño de la investigación que recoge la definición del contexto, los criterios de selección del caso de estudio, la metodología que

Capítulo 3. Metodología de la investigación

se seguirá para la recolección de datos y por último, un apartado sobre la fiabilidad y la validez de esta investigación.

3.1 Método de investigación

3.1.1 Enfoques de investigación

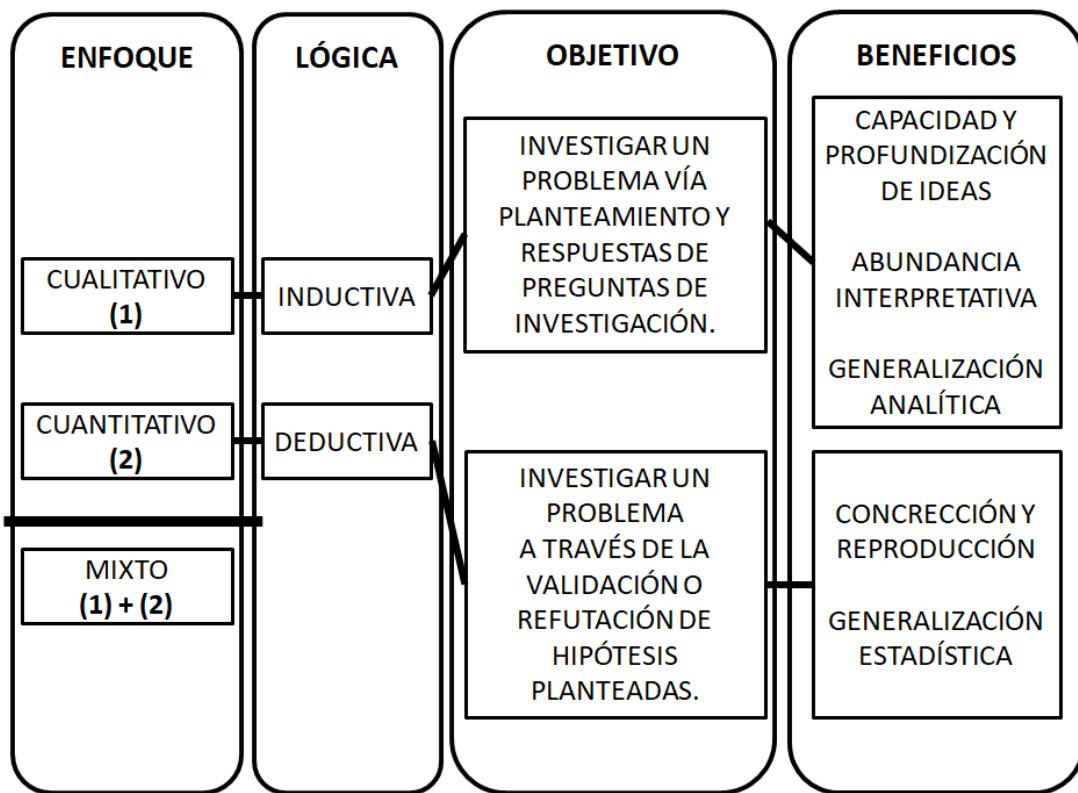
A día de hoy, existen tres diferentes enfoques en una investigación: el cualitativo, el cuantitativo y el mixto.

A continuación, se va a representar un esquema de resumen (Figura 3.1) que recoge de manera clara y sencilla las diferencias.

- **Enfoque Cualitativo:** este tipo de enfoque propone investigar sobre el problema, mediante el planteamiento de las preguntas de investigación y las respuestas a las mismas. Estas respuestas se obtendrán mediante la exploración y descripción de los respectivos casos que compongan la investigación, para luego generar teorías o modelos. Este tipo de enfoque sigue en la mayoría de los casos una lógica inductiva (Hernández Sampieri et al., 2006), es decir, el método científico es inductivo (Gray, 2013).
- **Enfoque Cuantitativo:** este tipo de enfoque propone investigar un problema concreto a través de la validación o refutación de las hipótesis planteadas. Dicha validación o refutación de hipótesis se llevará a cabo mediante técnicas estadísticas, siguiendo en la mayoría de los casos una lógica deductiva (Hernández Sampieri et al., 2006), es decir el método científico es deductivo (Gray, 2013).
- **Enfoque Mixto:** este tipo de enfoque es una combinación de los dos anteriores, basados en la combinación de los métodos inductivos y deductivos, el cual supone el ciclo completo de una investigación (Hernández Sampieri et al., 2006).

3.1 Método de investigación

Figura 3.1: Esquema de resumen de comparativa de enfoques de investigación



Fuente: Elaboración propia basada en (Hernández Sampieri et al., 2006; Gray, 2013)

3.1.2 Propuesta de investigación

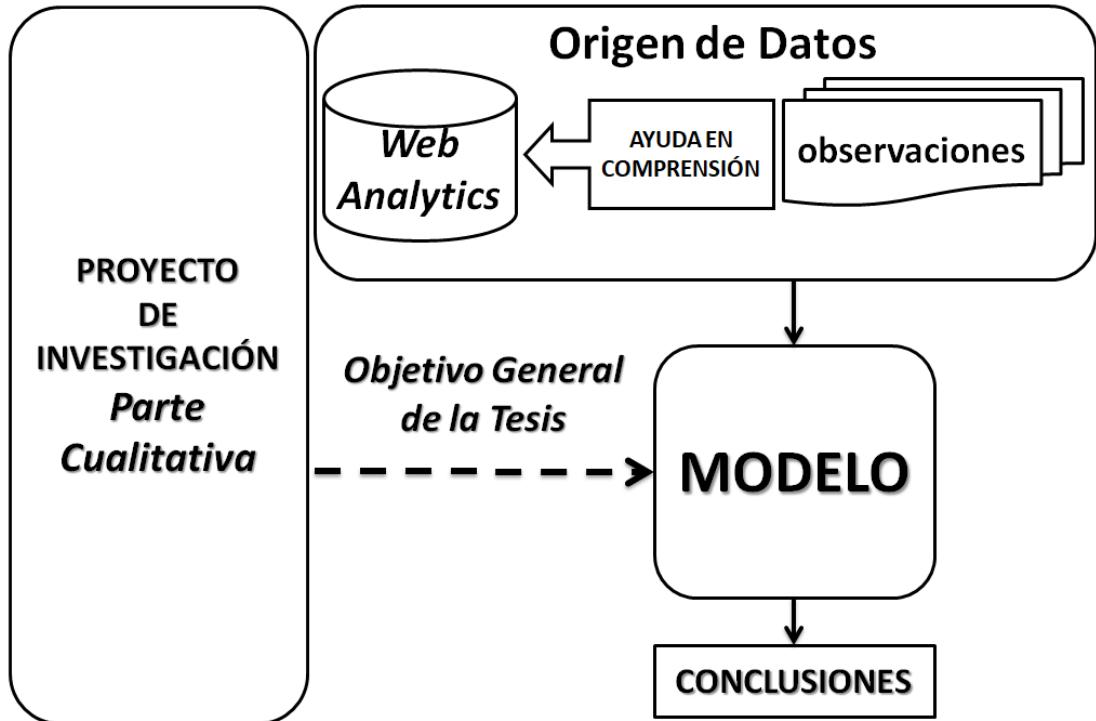
Teniendo en cuenta los diferentes tipos de enfoques, se ha decidido que el **enfoque cualitativo** es el que más se adapta al propósito de esta investigación. El motivo principal es que este tipo de enfoque, a través del planteamiento de preguntas de investigación, va a permitir descubrir, comprender e interpretar el fenómeno que se está investigando sobre "El comportamiento del visitante virtual".

Como consecuencia, se generará conocimiento sobre el comportamiento de los visitantes virtuales, que quedará recogido en el modelo que se presentará como resultado final de la investigación. De este modo, la metodología se ajusta al objetivo general del proyecto de investigación descrito en el punto de introducción. A

Capítulo 3. Metodología de la investigación

continuación, se muestra de manera gráfica lo expuesto hasta este punto, queriendo facilitar al lector la comprensión de la misma (véase Figura 3.2):

Figura 3.2: Esquema general de la metodología



Fuente: Elaboración propia

Dicho esto, y poniendo el foco en la parte cualitativa del esquema, a continuación se va a tratar el propósito de la investigación. Para ello, se va a tener en cuenta lo que afirma Castro (2010) sobre los tipos de estudios de casos. Se puede resumir la clasificación de estudios de casos propuesta por Yin (1994) en:

- Descriptivos, cuyo objetivo es analizar cómo ocurre un fenómeno organizativo dentro de su contexto real.
- Exploratorios, que buscan familiarizarse con un fenómeno o una situación sobre la que no existe un marco teórico bien definido.
- Ilustrativos, que de una u otra manera, ponen de manifiesto las prácticas de gestión de las empresas más competitivas.
- Explicativos, que tratan de desarrollar o depurar teorías, por lo que revelan las causas y los procesos de un determinado fenómeno organizativo (p.38)

3.1 Método de investigación

Tras observar la clasificación que planteaba Castro (2010), en base a lo propuesto por Yin (1994), es necesario indicar que, a nivel general, el propósito de la investigación se va a basar en un estudio de tipo **exploratorio**. Es decir, un estudio que busque explorar qué está ocurriendo y responder a las cuestiones que surjan alrededor de tal paradigma, a diferencia de otros como el descriptivo y el explicativo (Carazo, 2006).

Además, como se puede ver en el propio esquema, el origen de datos que servirá de fuente de información al modelo a generar es híbrido, es decir obtenidos a través de diferentes métodos de recolección de datos. En este sentido es importante recalcar que la parte del origen de datos asociada a la información obtenida vía observaciones, ayuda a comprender la parte de *Web Analytics*.

Una vez decidido que el enfoque cualitativo exploratorio es el más adecuado para esta investigación se tiene que decidir cual es el tipo de estudio cualitativo que mejor se adapta a los intereses perseguidos en este proyecto de investigación. Para ello, a continuación se muestra una comparativa realizada por Yin (1994), de cara a poder seleccionar cual es el tipo de estudio más adecuado para el proyecto de investigación.

Figura 3.3: Análisis de diversas estrategias para una investigación cualitativa

Estrategia o diseño	Preguntas de investigación
Experimento	¿Cómo? ¿Cuánto? ¿Por qué?
Encuestas	¿Quién? ¿Qué? ¿Dónde? ¿Cuánto?
Estudios históricos	¿Cómo? ¿Dónde? ¿Por qué?
Ánalisis de archivos	¿Quién? ¿Qué? ¿Dónde? ¿Cuánto?
Estudios de caso	¿Cómo? ¿Por qué?

Fuente: Yin (1994)

Basándose en Yin (1994) Figura 3.3, este proyecto se va a fundamentar en un **estudio de casos**. La razón principal por la cual se ha escogido, es porque se trata de estudiar el fenómeno dentro de su contexto, el comportamiento del visitante

Capítulo 3. Metodología de la investigación

virtual, en base a la información proporcionada por los participantes en la investigación tanto a nivel de *Web Analytics* como de observaciones.

Por otra parte, está la elección del tipo de "estudio de casos" a realizar. Las opciones son las siguientes:

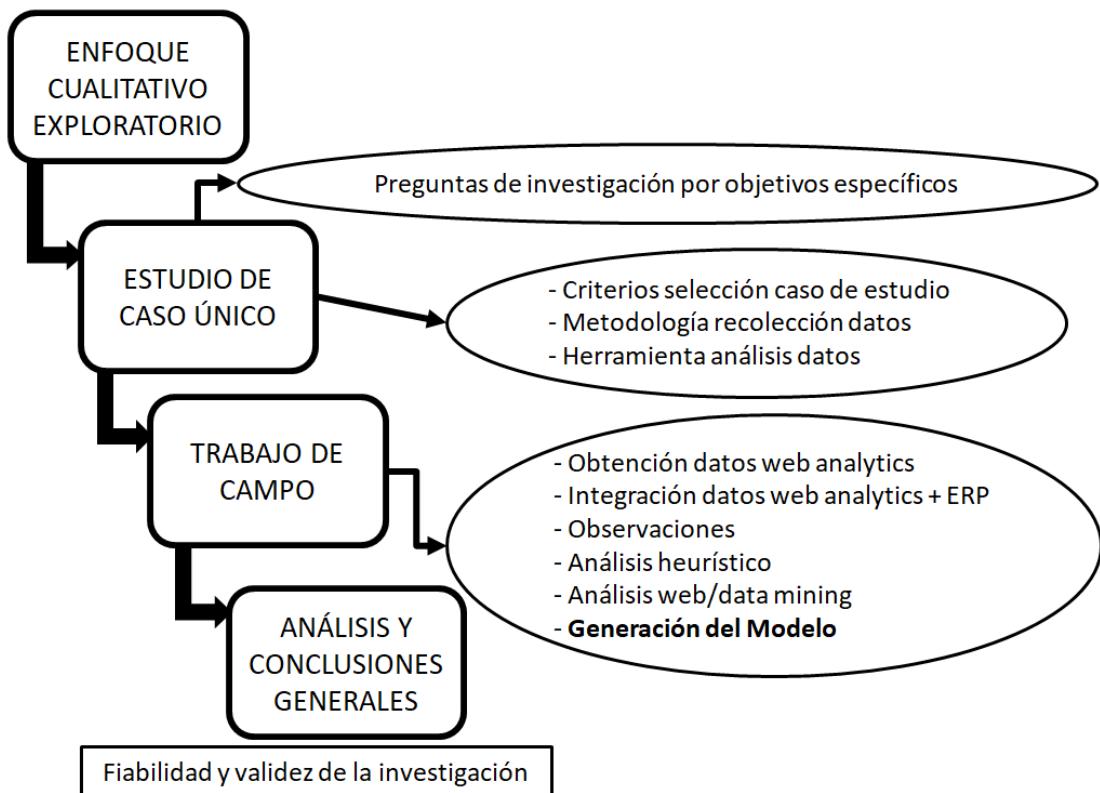
- **Caso único:** en este tipo de estudio de casos, el tratamiento del material del caso es suficientemente genérico y la calidad y la naturaleza de las conclusiones pueden llegar a ser únicas o fuertes, tal y como estableció Pettigrew (1990). Y por otro lado, el poseer todas las condiciones necesarias para confirmar, desafiar o ampliar una determinada teoría, hace que el caso sea especial, tal y como estableció Chiva Gómez (2001).
- **Casos múltiples:** la metodología basada en casos múltiples para Yin (1994) significaba el predominio del sentido de la réplica. Es decir, cada caso debería ser meticulosamente escogido y que cumpliese los siguientes requerimientos: a) en condiciones similares que generase resultados análogos a otro y b) por motivos predecibles, que concibiese resultados contrapuestos a otros casos. Otro autores como Ghauri & Grønhaug (2005) explicaron que, este tipo de estudio comparaba las respuestas de los diferentes casos para obtener conclusiones, en base a preguntar lo mismo a los diferentes casos. Añadiendo que la elección de caso caso debería de ser con un objetivo concreto. Es decir, evitando la selección por razón estadística y si aplicando motivos teóricos, ajustándose al paradigma que se quiere estudiar.

En este proyecto de investigación, se ha considerado que el **estudio de casos** que se va a llevar a cabo va a ser el de **caso único**, ya que éste ofrece evidencias sólidas a diferencia del caso múltiple, en el cual se producen evidencias parciales generadas por cada caso. En este sentido, el hecho de que la mayoría de las organizaciones identificadas no dispongan de todo el sistema de comercio electrónico y además no lo tengan integrado con la parte de su sitio web corporativo, tal y como se menciona en el punto 3.2.3, ha inclinado la balanza a favor de disponer de un solo caso de estudio sólido.

3.2 Diseño de la investigación

En este apartado, se presenta el diseño de la investigación que se va a llevar a cabo. Incluye la definición del contexto, así como la selección del caso de uso, la metodología aplicada en la recolección de los datos, la herramienta de análisis de datos y por último, el trabajo de campo. A continuación se va a exponer un esquema con el propósito de facilitar la comprensión y el hilo conductor que va a seguir la propuesta metodológica.

Figura 3.4: Esquema de resumen de la propuesta de metodología



Fuente: Elaboración propia basada en (Monge, 2010)

3.2.1 Preguntas de investigación

Este apartado tiene el propósito de describir las preguntas de investigación de esta tesis doctoral teniendo en cuenta la trazabilidad con sus objetivos específicos correspondientes.

Capítulo 3. Metodología de la investigación

3.2.1.1 Preguntas asociadas al primer objetivo específico

El primer bloque de preguntas pone el foco en identificar los componentes de la cadena de suministro de ORBEA S.Coop., analizar la evolución de la cadena física a la virtual, y más concretamente, del cliente/visitante físico al virtual.

Objetivo 1: conocer la evolución de la cadena de suministro desde la configuración física hasta la virtual.

Para ello se plantean las siguientes preguntas de investigación:

P1.1.- ¿Cuáles son los componentes de la cadena de suministro y la relación entre ellos?

P1.2.- ¿Cuál es la evolución tanto de la cadena como del cliente físico al virtual?

3.2.1.2 Preguntas asociadas al segundo objetivo específico

Este bloque de preguntas se centra en la cadena de suministro virtual del caso de estudio. Se trata de identificar los escenarios de relación de ORBEA S.Coop. con los diferentes tipos de clientes/visitantes, las ramas de análisis de comportamiento asociadas y así, analizar el comportamiento del visitante virtual. En este sentido, es importante recordar que una rama de análisis de comportamiento es un ámbito del caso de estudio, que abarca un tema determinado asociado a los escenarios de relación, tal y como se comenta en el punto 2.2.2.

Objetivo 2: analizar el impacto que tiene en la propia organización, el comportamiento del visitante virtual en los escenarios de relación de la cadena de suministro.

Para ello se plantean las siguientes preguntas de investigación:

P2.1.- ¿Cuáles son los diferentes escenarios de relación de ORBEA S.Coop. con los clientes/visitantes virtuales?

P2.2.- ¿Cuáles son las ramas de análisis de comportamiento del visitante virtual asociadas a cada escenario de relación?

3.2 Diseño de la investigación

P2.3.- ¿Qué conclusiones se pueden obtener de la observación del comportamiento de los visitantes virtuales de forma aislada y comparativamente con las ventas realizadas?

3.2.1.3 Preguntas asociadas al tercer objetivo específico

En este bloque se pone el foco en las ramas de análisis de comportamiento asociadas a los diferentes escenarios de relación que tendrá el caso de estudio, para que se den respuesta mediante el análisis de los datos a las preguntas de investigación asociadas.

Objetivo 3: identificar los factores que permiten aproximar a los proveedores de primer nivel, al resto de agentes de la cadena de suministro a nivel de comportamiento del visitante virtual.

Para ello se plantean las siguientes preguntas de investigación:

P3.1.- ¿Cuáles de los valores de las características de las tecnologías de la información (IT) se tendrán en consideración en los escenarios de relación?

P3.2.- ¿Cuáles son las relaciones, tanto entre los tipos, como en el acceso a los productos visitados, en la web y los vendidos?

3.2.1.4 Preguntas asociadas al cuarto objetivo específico

En este bloque se pone el foco en poder definir un conjunto de pautas, que permitan generar una estrategia analítica en beneficio del caso de estudio del proyecto de investigación. De esta manera, se logrará aplicar el conocimiento que se obtenga del comportamiento del visitante virtual, en el desarrollo del sitio web y otros desarrollos tecnológicos futuros.

Objetivo 4: definir la estrategia analítica para aplicar el conocimiento sobre el comportamiento del visitante virtual, en el desarrollo del sitio web y otros desarrollos tecnológicos.

Para ello se plantea la siguiente pregunta de investigación:

P4.1.- ¿Cuáles son las pautas para construir una estrategia analítica?

Capítulo 3. Metodología de la investigación

3.2.2 Definición del contexto

La investigación se va a centrar en un caso de estudio, perteneciente a los proveedores de primer nivel del sector de distribución en España. En este sentido, es importante recalcar dos aspectos: por un lado, la importancia del sector de la distribución y gran consumo y el establecimiento del foco en los proveedores de primer nivel, debido a que:

1. Según AECOC, en España el sector representa el 20 % del PIB (Bonmatí, 2015).
2. El sector canaliza, en mayor o menor medida, el desarrollo del comercio electrónico, a través del recorrido que están teniendo las plataformas de comercio electrónico.
3. El comercio electrónico supone un punto de encuentro de clientes y no clientes, en definitiva, visitantes virtuales de los cuales, comprender su comportamiento resulta fundamental.
4. A medida que en la cadena de valor del sector de la distribución se han ido incorporando los canales de venta online, el análisis del comportamiento del visitante virtual se ha orientado de manera clara hacia los consumidores finales y hacia los elementos de distribución (distribuidores y gran consumo).
5. Los trabajos revisados tienen un origen de datos mayoritariamente basado en encuestas siendo completados en su mayoría con análisis estadísticos.
6. A pesar del sesgo comentado hacia los consumidores y los distribuidores, no hay que olvidar que en determinados casos no está claro a qué participantes de la cadena de valor corresponden las relaciones cliente-proveedor.

Por otro lado, dentro del sector de distribución y concretamente en los proveedores de primer nivel de bicicletas, ORBEA S.Coop. es una empresa de referencia tanto a nivel nacional como a nivel internacional (Iturbe-Ormaetxe, 2018), recordando que es un caso de estudio sólido, ya que abarca todas las casuísticas posibles, tal y como se analiza en el siguiente apartado.

3.2.3 Criterios de selección del caso de estudio

Una vez definido el contexto y tras decidir, tal y como se comenta en el apartado anterior, que el proyecto de investigación se va a basar en un único caso, en este apartado se describe el proceso seguido hasta llegar a la elección del mismo. Los pasos seguidos se enumeran a continuación:

1. Elaboración de una lista inicial de empresas del sector de distribución en España. A partir de la información proporcionada por las asociaciones como AECOC, AFIDAD (Asociación Española de Distribución Deportiva), ANGED (Asociación Nacional Grandes de Empresas de Distribución) y AMBE (Asociación de Marcas y Bicicletas de España) y el portal web "ElEconómista.es".
2. Selección de la lista inicial, por el criterio de empresas que son proveedoras de primer nivel.
3. Filtrado dependiendo de si disponen de B2B, B2C y web corporativa con catálogo.
4. Selección según si tienen integrado el B2B y B2C y el catálogo web bajo un mismo sitio web corporativo.
5. Filtrado por el criterio de accesibilidad: su interés en colaborar y su grado de accesibilidad directa en el proyecto de investigación.

Tras aplicar los filtros citados anteriormente y tras valorar la importancia de ORBEA S. Coop. como proveedora de primer nivel dentro del sector de la distribución, se optó por su selección como caso único del proyecto de investigación.

3.2.4 Metodología de recolección de datos

La metodología de recolección de datos que va a seguir esta tesis doctoral tiene dos partes: en primer lugar, el conjunto de pasos que se va a realizar con cada fuente del origen de datos y segundo lugar, cada pregunta de investigación con qué fuente va a ser respondida, con el propósito de enfocar la recogida de datos a las necesidades de cada pregunta de investigación.

Capítulo 3. Metodología de la investigación

3.2.4.1 Origen de datos

El origen de datos para este proyecto de investigación proviene de dos fuentes. Por un lado, la recogida de datos cuantitativos mediante herramientas de *Web Tracking* y, por otro lado, las observaciones del fenómeno a investigar. Para poner ambas en funcionamiento, se van a seguir los siguientes pasos:

Tabla 3.1 Proceso del Origen de Datos

	Procesos
Origen de Datos <i>Web Analytics</i>	<ol style="list-style-type: none">1. Establecimiento del sistema de <i>Web Tracking</i> en el sitio web.2. Acuerdo con el caso de estudio sobre la accesibilidad de los datos (cuáles se van a obtener, durante cuánto tiempo y quién es titular de los datos, etc.).3. Por medio de soluciones tecnológicas de BI (<i>Business Intelligence</i>), se realizará la integración de los datos en una base o almacén de datos. Todo ello, a través de procesos de ETL (<i>Extraction, Transformation and Load</i>) que se describen en detalle en el siguiente capítulo.
Observaciones	<ol style="list-style-type: none">1. Realización de observaciones respecto al caso de estudio para ayudar a comprender la parte de <i>Web Analytics</i>2. Recogida de las observaciones tras los diferentes encuentros que se tendrán con el caso de estudio.3. Registro de la información que se recogerá en las observaciones en documentos o en otro tipo de canal digital (correo electrónico, sistema de ficheros de <i>cloud</i> como <i>Dropbox</i>, etc.)

Fuente: Elaboración propia

Una vez visto el proceso anterior, es necesario que cada pregunta de investigación esté alineada, tanto con su objetivo específico como con las fuentes del origen

3.2 Diseño de la investigación

de datos correspondientes. De esta forma, se obtiene una trazabilidad entre las preguntas y la fuente necesaria para que puedan ser respondidas, tal y como se puede ver en la siguiente tabla.

Tabla 3.2 Origen de datos que da respuesta a cada pregunta de investigación

Objetivo específico	Pregunta	Fuente para la respuesta
1	P1.1	
	P1.2	
	P2.1	Observaciones del fenómeno
2	P2.2	
	P2.3	Datos cuantitativos recogidos mediante <i>Web Analytics</i> y Observaciones del fenómeno
3	P3.1	
	P3.2	Datos cuantitativos recogidos mediante <i>Web Analytics</i>
4	P4.1	Datos cuantitativos recogidos mediante <i>Web Analytics</i> y Observaciones del fenómeno

Fuente: Elaboración propia

3.2.4.2 Selección de las herramientas de *Web Tracking*

Dentro del proceso de *Web Analytics* está la etapa de obtención de los datos que se van generando, debido a las visitas que realizan los visitantes/clientes virtuales en el sitio web, lo que se denomina *Web Tracking*. A continuación se va a explicar que herramientas de *web tracking* existentes a día de hoy en el mercado se van a utilizar en la propia investigación.

De las posibles herramientas existentes (ver apartado 2.3.1, tablas 2.4 y 2.5) se ha utilizado *Google Analytics* y *Piwik*. Los motivos fueron que *Google Analytics* es la herramienta propietaria que mejor características presentaba en los estudios y además el caso de estudio era favorable a su utilización, debido al conocimiento sobre la misma. En el caso de *Piwik* entre las herramientas de *open source* era la herramienta más favorable (similitud con *Google Analytics*, tracking de diferentes sitios web, etc.) y también por la opción que ofrecía de poder disponer de una infraestructura *IT* en la nube a través de un software como servicio. Sin embargo, se tuvo que cancelar la utilización de esta última herramienta, debido a que la empresa

Capítulo 3. Metodología de la investigación

que nos facilitaba el servicio no quiso mantener las condiciones iniciales orientadas a un proyecto de investigación (véase el capítulo 6).

Por todo ello, *Google Analytics* es la herramienta de web tracking que se va a utilizar en el *Web Analytics*, para este proyecto de investigación.

3.2.5 Herramienta de análisis de datos

La herramienta de análisis de datos a utilizar en este proyecto de investigación, va a ser la *suite* de BI (*Business Intelligence*): *MicroStrategy*. Debido sobre todo a la capacidad integradora de esta herramienta, permite por ejemplo disponer del *software R* sin necesidad de tener un servidor dedicado. Además, en el lenguaje de programación estadístico y de análisis, R, se dispone de múltiples ejemplos con las técnicas de *Web Mining* descritas en la revisión de la literatura.

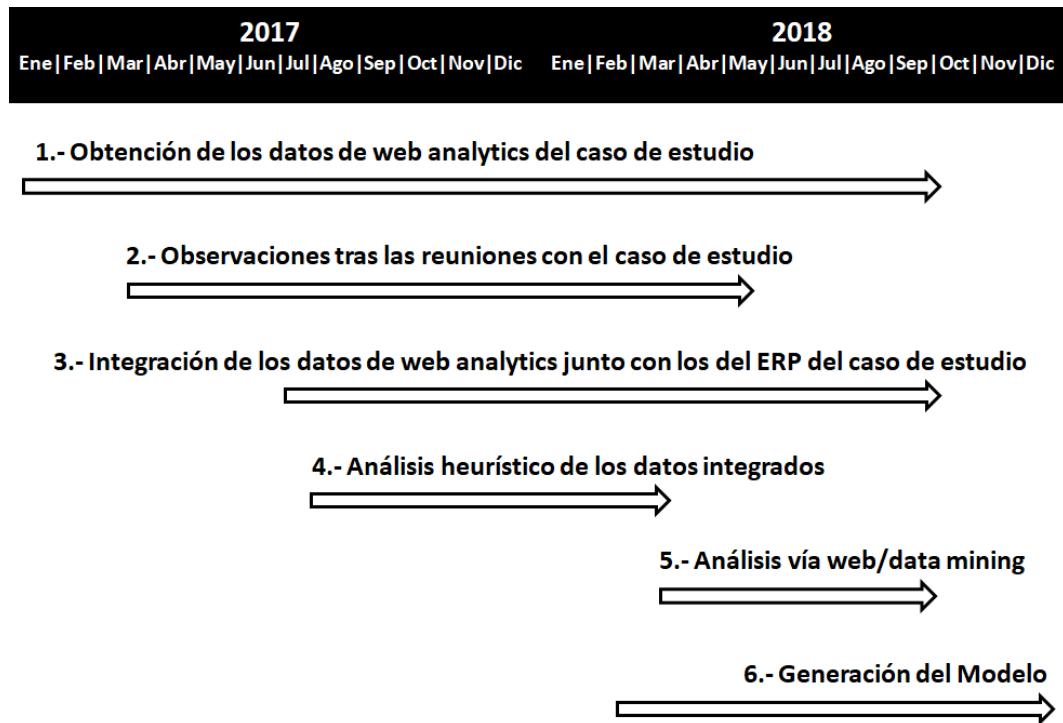
Asimismo, en determinadas funcionalidades del propio *MicroStrategy* se proveen de técnicas de *Web Mining* (véase capítulo 2) como por ejemplo la predicción a través de líneas de tendencias.

3.2.6 Trabajo de campo

El trabajo de campo realizado en esta investigación abarcó el periodo desde Enero de 2017 hasta Diciembre de 2018. A continuación se muestra un esquema del mismo:

3.2 Diseño de la investigación

Figura 3.5: Cronograma de trabajo de campo



Fuente: Elaboración propia

Tal y como se puede observar en la figura anterior, existen dos bloques principales dentro del trabajo de campo: *Web Analytics* y *Observaciones*. Estas se llevaron a cabo a través de las siguientes tareas:

- *Web Analytics*:

1. **Obtención de los datos de *Web Analytics* del caso de estudio:** tarea que ha facilitado obtener los datos del comportamiento de los visitantes/clientes virtuales en el sitio web del caso de estudio, a través de la utilización de *Google Analytics*.
2. **Integración de los datos de *Web Tracking* con los del ERP del caso de estudio:** tarea que ha permitido almacenar los datos provenientes de *Google Analytics*, junto con los datos del ERP del caso de estudio. Todo ello, mediante el desarrollo y posterior ejecución de *software* que implementa procesos de ETL, que se describirán en el siguiente capítulo.

Capítulo 3. Metodología de la investigación

Para este proyecto se dispone de un conjunto de datos desde Enero del 2017 hasta Septiembre del 2018.

3. **Análisis heurístico de los datos integrados:** tarea que ha facilitado el análisis y comparativa de los datos a través de la herramienta MicroStrategy (véase el punto 3.2.5).
 4. **Análisis vía Web/Data Mining:** tarea que ha permitido a través del *software* de BI: *Microstrategy*, analizar a través de la técnica de predicción, disponible en la propia herramienta mediante las líneas de tendencia, los datos integrados anteriormente.
 5. **Generación del modelo:** tarea que ha llevado a cabo un esquema que reúne las respuestas obtenidas en las diferentes preguntas de investigación y concretamente los análisis indicados anteriormente.
- Observaciones: tarea mediante la cual, tras las reuniones con el caso de estudio, se ha obtenido conocimiento sobre el propio caso, para poder responder las preguntas de investigación correspondientes y también para complementar la información obtenida del *Web Analytics*.

3.3 Fiabilidad y validez de esta investigación

Uno de los aspectos claves de una investigación es certificar la fiabilidad y la validez interna y externa de la misma. Según afirma Castro (2010) "Sutton (1997) señala que, para que una investigación sea considerada valiosa y aceptable, tiene que ser rigurosa, por lo que debe ser fiable, válida y generalizable.". Asimismo, según Franklin (2005) la fiabilidad es "el grado en que diferentes investigadores, que recolecten datos similares en un mismo campo y efectúen los mismos análisis, generen resultados equivalentes". Es decir, si otro investigador/a estudiase la información que se ha adquirido, éste/a debería llegar a unas conclusiones coherentes con las obtenidas en esta investigación.

La validez interna se refiere a la credibilidad de la investigación, es decir, si el investigador ha comprendido la información proporcionada por los participantes y más concretamente, todo lo relacionado con el tema de estudio. Según Mertens

3.3 Fiabilidad y validez de esta investigación

(2014) la credibilidad es "la relación entre la forma en que el participante percibe los conceptos vinculados al tema planteado y la forma en que el investigador retrata los puntos de vista de cada uno de ellos".

La validez externa se refiere a la aplicabilidad de los resultados al finalizar el proyecto de investigación. No se trata de generalizar los resultados a una población más amplia, sino de ver qué parte de estos resultados pueden aplicarse en otros contextos (Williams, Tutty, & Grinnell, 2005). En el caso de este proyecto de investigación, no se pretende una generalización estadística o inferencia, sino que se realizará una **generalización analítica**, ya que tal y como dice Carazo (2006) esta última supone: "utilizar el estudio de caso único o múltiple para ilustrar, representar o generalizar a una teoría". Incluso, otro autor importante como es Yin, también respalda la generalización analítica cuando el tipo de estudio es exploratorio (Yin, 1994). De hecho, Carazo (2006) afirma que:

La cuestión de la generalización de los estudios cualitativos (incluido el estudio de caso) no radica en una muestra probabilística extraída de una población a la que se pueda extender los resultados, sino en el desarrollo de una teoría que puede ser transferida a otros casos. De aquí que algunos autores prefieran hablar de transferibilidad, en vez de generalización, en la investigación de naturaleza cualitativa (Maxwell, 1998). (p.173)

En este sentido, el concepto de transferibilidad se tiene en cuenta de cara a la validez y fiabilidad de esta investigación (véase el punto 3.3) y concretamente en la validez externa de la investigación.

Por consiguiente, en el desarrollo del proyecto de investigación se perseguirá el cumplimiento de ambos criterios de validez y fiabilidad, llevando a cabo acciones que aseguren el cumplimiento de los mismos. Aparte de la validez y fiabilidad de la propia investigación, es necesario garantizar el cumplimiento de la Ley de Protección de Datos y la confidencialidad de los datos proporcionados, por ser relevantes para el desarrollo del negocio. Y además, que dichos datos no puedan llegar a manos de la competencia, que pudiera hacer un uso incorrecto de ellos y poner en peligro la ventaja competitiva que supone el conocimiento obtenido.

Capítulo 3. Metodología de la investigación

En definitiva, se tomarán un conjunto de medidas para garantizar la transparencia y el respeto a la propiedad de los datos utilizados en la investigación.

Si los demás pueden tú también.

Anónimo

CAPÍTULO

4

Análisis de la evidencia empírica

Disclaimer: el contenido de este capítulo no está disponible por cuestiones de derechos intelectuales/industriales.

Algunas personas quieren que suceda, otras desean que suceda, otras hacen que ocurran.

Michael Jordan

CAPÍTULO
5

Generación del modelo

Disclaimer: el contenido de este capítulo no está disponible por cuestiones de derechos intelectuales/industriales.

Imaginemos el futuro y vayamos hacia él. No dejemos que el tiempo nos arrastre.

Walt Disney

CAPÍTULO
6

Conclusiones, limitaciones y líneas futuras

Una vez que el proyecto ha concluido, a continuación, se van a exponer las conclusiones, limitaciones y líneas futuras. Para ello, en primer lugar, se va a realizar una revisión de los objetivos y las preguntas de investigación asociadas, comentando los resultados y las respuestas obtenidas a las propias preguntas. A continuación, se van a presentar las conclusiones del proyecto e inmediatamente después, se van a mostrar las aportaciones que realiza este proyecto en beneficio tanto para el ámbito científico-académico como para el ámbito empresarial. Finalmente, se van a explicar las limitaciones y líneas futuras identificadas.

6.1 Revisión de los objetivos

En este apartado, se va realizar una recapitulación de los objetivos específicos del proyecto y también se va exponer el grado de cumplimiento de cada uno de ellos, en base al siguiente criterio: el número de preguntas de investigación asociadas con respuesta y a cómo de completas han sido las respectivas respuestas. Los objetivos específicos son los siguientes:

Capítulo 6. Conclusiones, limitaciones y líneas futuras

- **Objetivo 1:** conocer la evolución de la cadena de suministro desde la configuración física hasta la virtual.
- **Objetivo 2:** analizar el impacto que tiene en la propia organización, el comportamiento del visitante virtual en los escenarios de relación de la cadena de suministro.
- **Objetivo 3:** identificar los factores que permiten aproximar a los proveedores de primer nivel, al resto de agentes de la cadena de suministro a nivel de comportamiento del visitante virtual.
- **Objetivo 4:** definir la estrategia analítica para aplicar el conocimiento sobre el comportamiento del visitante virtual, en el desarrollo del sitio web y otros desarrollos tecnológicos.

Entrando en detalle:

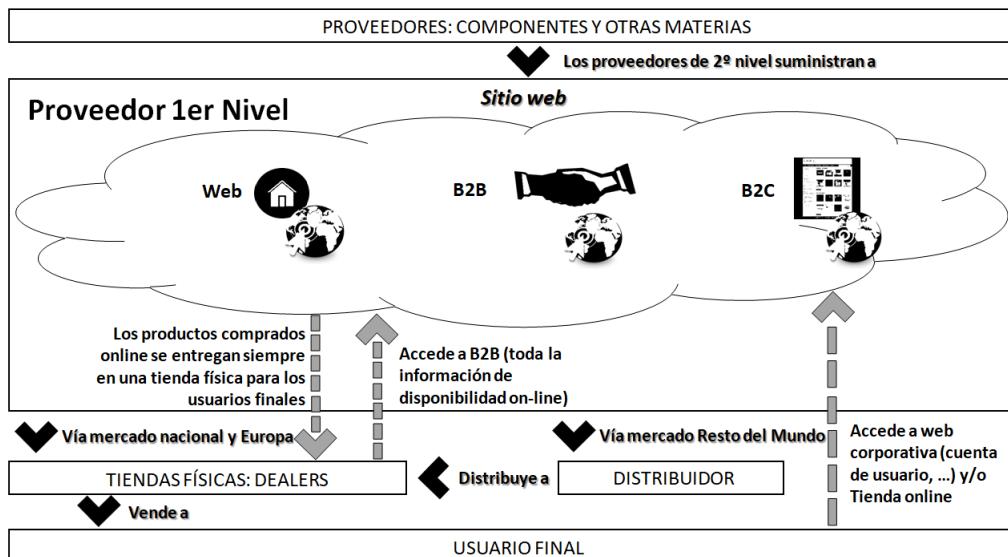
- **Objetivo 1:**

Conocer la evolución de la cadena de suministro desde la configuración física hasta la virtual.

Resultado: se ha conseguido conocer la evolución de la cadena de suministro desde la configuración física hasta la virtual de un proveedor de primer nivel y así poder disponer de una estructura de base (Figura 6.1) para construir el modelo:

6.1 Revisión de los objetivos

Figura 6.1: Propuesta de estructura global de la cadena de suministro física y virtual (Recapitulación)



Fuente: Elaboración propia

Grado de cumplimiento: para el objetivo 1 es del 100 %. Debido a que se han conseguido responder las dos preguntas de investigación asociadas, tal y como se verá en el siguiente apartado. Con las respuestas de ambas, se ha generado una cadena de suministro física y virtual, a tener en cuenta, para los proveedores de primer nivel del sector de distribución. Todo ello ha permitido establecer el papel de los diversos agentes que interceden en la misma.

- **Objetivo 2:**

Analizar el impacto que tiene en la propia organización, el comportamiento del visitante virtual en los escenarios de relación de la cadena de suministro.

Resultado: se ha logrado analizar el impacto que tiene en la propia organización, el comportamiento del visitante virtual en los escenarios de relación de la cadena de suministro. Todo gracias a la identificación de las ramas de análisis de comportamiento del visitante virtual que están asociados a cada escenario de relación. Sin olvidar, que este caso de estudio dispone de todos los escenarios de relación identificados. Por todo ello, el análisis del impacto se ve reflejado en lo

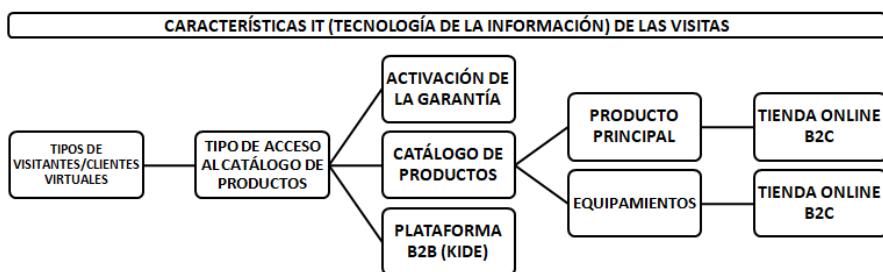
Capítulo 6. Conclusiones, limitaciones y líneas futuras

siguiente.

Por un lado, del esquema del flujo de la visita del visitante virtual (Figura 6.2), en el cual se reflejan en las diversas entidades, las ramas de análisis definidas y los propios escenarios de relación.

Por otro lado, en el impacto que produce el cruce de cada rama de análisis de comportamiento con la información proveniente del *ERP* del caso de estudio.

Figura 6.2: Flujo de la visita de los visitantes virtuales (Recapitulación)



Fuente: Elaboración propia

Grado de cumplimiento: para el objetivo 2 es del 90 %. Debido a que se han conseguido responder de manera completa a dos de las tres preguntas de investigación asociadas. Sin embargo, en el caso de la tercera, no ha sido del todo completa ya que una de las ramas, CVD (Captación y Vigilancia Digital), no ha tenido el desarrollo esperado en el proyecto, tal y como se va a explicar en las líneas futuras.

• **Objetivo 3:**

Identificar los factores que permiten aproximar a los proveedores de primer nivel, al resto de agentes de la cadena de suministro a nivel de comportamiento del visitante virtual.

Resultado: se han identificado los diferentes factores y cada uno que recorriendo tiene, tal y como se muestra en la Figuras que se presentan a continuación. De esta forma se permite aproximar a los proveedores de primer nivel, al resto de agentes de la cadena de suministro a nivel de comportamiento del visitante

6.1 Revisión de los objetivos

virtual. Todos los factores que se detallan a continuación, van a aportar conocimiento al proveedor de primer nivel. En total un conjunto de cinco factores: factor Equilibrio CDD (Control del Canal Digital del Distribuidor), factor Madurez CAP (Control de tipo de Acceso al Producto), factor Plataforma ECO (Evolución de la Conexión), factor Proporción CVD y factor Tendencia REP (Recambios de Productos). Para medir cada uno de ellos, se ha utilizado una escala ordinal (bajo, medio, alto), que ofrece una gama simple y sencilla de valores, fáciles de comprender para los agentes del caso de estudio. A continuación, se muestran cada uno de ellos.

Primero, está el factor denominado Equilibrio CDD, para que el proveedor de primer nivel conozca cómo de acorde es el interés *online* de sus visitantes virtuales, con lo que compran sus clientes. Este factor puede tener tres posibles valores: nivel bajo, medio o alto. Todos ellos se obtienen en base a las condiciones que se muestran en el siguiente esquema, las cuáles, han sido identificadas tras el análisis de la evidencia empírica (véase capítulo 4) realizado.

Después, está el factor llamado madurez del CAP, el cual sirve al proveedor de primer nivel para conocer en qué nivel desarrollo está el acceso al tipo de producto en su sitio web, respecto a los tipos de visitantes virtuales expertos y nóveles que acceden al propio sitio. Este factor va a tener un valor de "Alto" cuando el número de accesos de los visitantes virtuales de tipo "Experto" sea mayor y con diferencia respecto al número de accesos de los visitantes virtuales de tipo "Novel" y viceversa. En cambio, si la diferencia en un sentido u en otro es insignificante o irrisoria entonces el factor obtendrá el valor "Medio".

En tercer lugar, está el factor denominado plataforma ECO. Este factor sirve para conocer el tipo de plataforma para los dos entornos: B2B y B2C. El tipo de plataforma nos da información sobre el sistema operativo (SO), navegador web y dispositivo con el que se accede. Este factor va a tener tres posibles niveles para los dos entornos: "plataforma escritorio", "mixta" y "móvil". El primer valor se asigna si los accesos de los visitantes virtuales van a ser en gran medida a través

Capítulo 6. Conclusiones, limitaciones y líneas futuras

de SO, navegador web y dispositivo de escritorio. En cambio, si los accesos fueran similares (entre plataforma de escritorio y móvil), se le asigna el valor de "plataforma mixta". Por último, se le asigna el valor "plataforma móvil" si los accesos son mayoritariamente a través del dispositivo de móvil.

El cuarto factor, llamado Proporción CVD, permite al proveedor de primer nivel conocer cuál es la correlación entre los visitantes virtuales corporativos y las tipologías de clientes de que dispone. Este factor puede tener tres posibles niveles: alto, medio y bajo. El primer valor se asigna si los accesos son en su mayoría de los visitantes virtuales corporativos que son clientes de tipo real para el propio proveedor de primer nivel. En cambio, si los accesos de los visitantes virtuales corporativos solo coinciden con determinado conjunto de clientes reales del proveedor, se le asigna el valor "Medio". Por último, se le asigna el valor "Bajo" si los accesos de los visitantes virtuales corporativos apenas coinciden con los clientes reales. Es decir, en este caso el factor demostraría que los clientes reales del proveedor no coinciden con los visitantes virtuales corporativos.

El último factor, es el llamado tendencia REP que sirve para conocer en qué proporción accederán los visitantes virtuales a los diferentes equipamientos que dispone el proveedor de primer nivel en la parte de recambios de productos. Este factor dispone de 3 valores: bajo, medio y alto. Siguiendo las mismas directrices en cuanto al tipo de escala, tal y como se muestra en la propia imagen.

Grado de cumplimiento: para el objetivo 3 es del 90 %. Debido a que se han conseguido responder las dos preguntas de investigación asociadas, pero el factor asociado a la rama CVD no ha tenido el desarrollo debido, ya que, esta rama (como se ha dicho anteriormente) no ha tenido el desarrollo esperado.

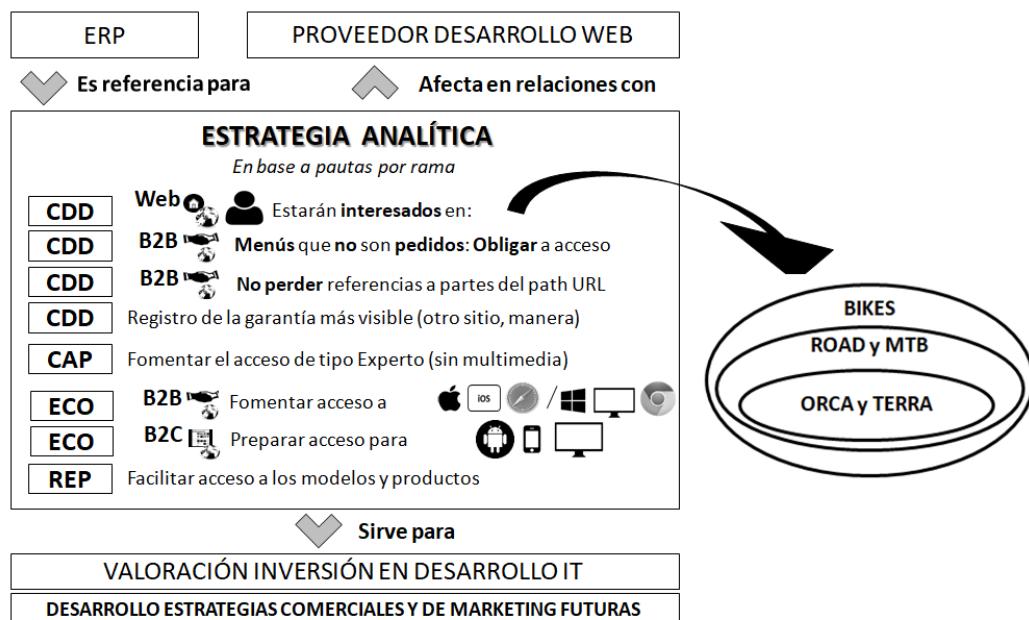
- **Objetivo 4:**

Definir la estrategia analítica para aplicar el conocimiento sobre el comportamiento del visitante virtual, en el desarrollo del sitio web y otros desarrollos tecnológicos.

6.2 Revisión de las preguntas de investigación

Resultado: se ha podido generar la estrategia analítica, para emplear lo aprendido sobre el comportamiento del visitante virtual, en la evolución del sitio web y otras implementaciones tecnológicas.

Figura 6.3: Esquema de la estrategia analítica (Recapitulación)



Fuente: Elaboración propia

Grado de cumplimiento: para el objetivo 4 es del 85 %. Debido a que se han conseguido responder la pregunta de investigación asociada, pero el factor asociado a la rama CVD no ha podido aportar ninguna pauta por las razones anteriormente mencionadas.

6.2 Revisión de las preguntas de investigación

En este apartado, se va realizar una recapitulación de las preguntas de investigación y de sus respectivas respuestas dadas en este documento.

P1.1.- ¿Cuáles son los componentes de la cadena de suministro y la relación entre ellos? La cadena de suministro del caso de estudio en su versión física,

Capítulo 6. Conclusiones, limitaciones y líneas futuras

empieza con lo referente a la cadena de montaje. Es decir, los proveedores de segundo nivel que proveen al propio caso de estudio, de los componentes necesarios para montar sus productos, y también otro tipo de materias necesarias para el proceso de fabricación. Tras finalizar el proceso de montaje, los productos tienen dos vías de distribución: 1) Nacional y Europa: las tiendas físicas (*Dealers*) dónde se realizan las ventas de los productos. Y 2) Resto del Mundo: el distribuidor que luego distribuye a las tiendas del país correspondiente.

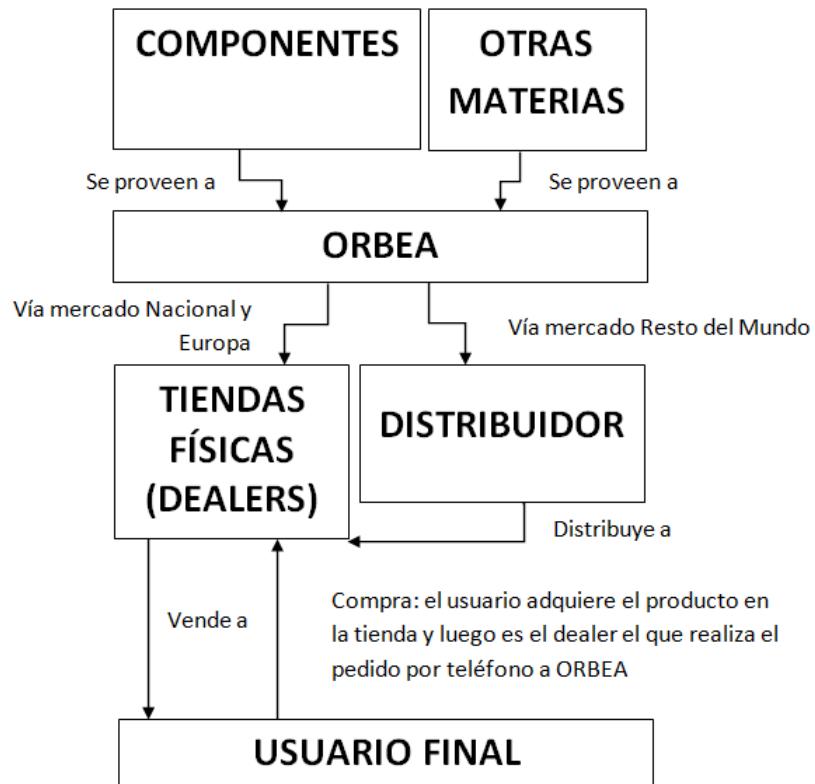
En lo referido a la compra-venta, entre la parte de distribución y el usuario final, en la versión física de la cadena, antes de la irrupción de Internet, el usuario final compraba los productos en las tiendas, dentro de lo que suponía una relación directa (*Brick and mortar*). Por lo que era la propia tienda quién hacía los pedidos por teléfono al caso de estudio. En consecuencia, no había una relación entre el proveedor de primer nivel y el usuario final.

P1.2.- ¿Cuál es la evolución tanto de la cadena como del cliente físico al virtual?

La evolución tanto de la cadena (Figuras 6.4 y 6.5) como del cliente físico al virtual es la que se muestra a continuación:

6.2 Revisión de las preguntas de investigación

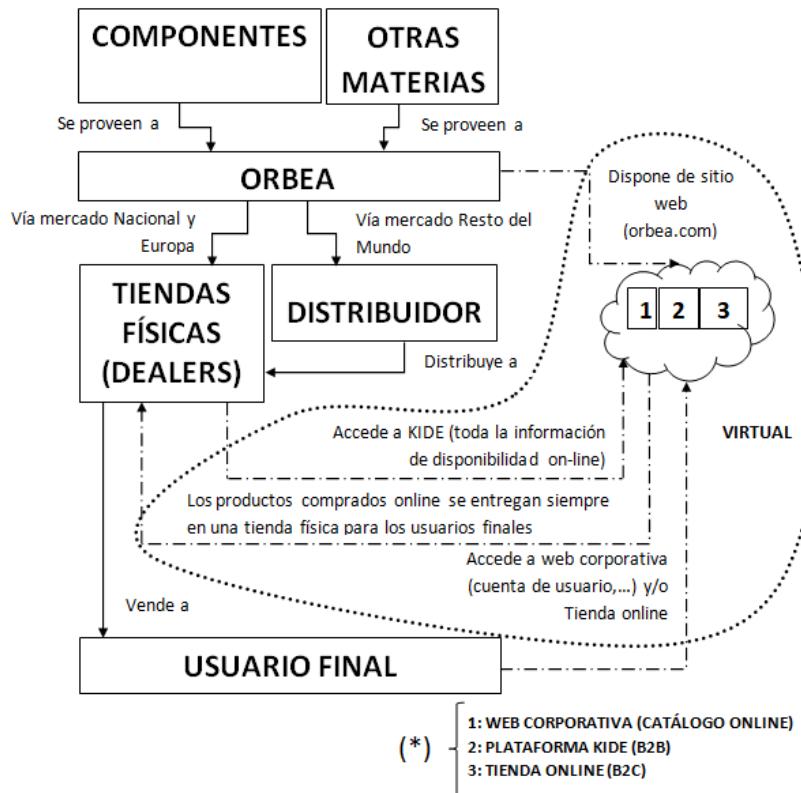
Figura 6.4: Cadena de suministro física (Recapitulación)



Fuente: Elaboración propia

Capítulo 6. Conclusiones, limitaciones y líneas futuras

Figura 6.5: Cadena de suministro física + virtual (Recapitulación)



Fuente: Elaboración propia

El concepto de cliente físico, ha ido evolucionando al concepto de visitante virtual, en la cadena de suministro de ORBEA S.Coop. Es decir, la llegada de los sistemas de comercio electrónico y su consiguiente integración en su sitio web, ha permitido, por un lado, llevar a cabo el mantenimiento del cliente físico tradicional: sus tiendas físicas, pero con el cambio de paradigma que supone pasar de realizar los pedidos por teléfono y relación directa, a acceder a toda la información de disponibilidad de los productos *online* y recibir los productos comprados *online*. Es decir, se convierte también en un visitante virtual de la plataforma B2B KIDE. Y, por otro lado, se ha podido llevar a cabo el acercamiento del usuario final al proveedor: a través del rol de visitante virtual (el propio ORBEA S.Coop.) a nivel de producto/servicio posibilitando el acceso a su cuenta de usuario y/o al catálogo de la tienda *online*. De esta forma puede comenzar el proceso de compra, evolucionando el rol de visitante virtual a cliente virtual.

6.2 Revisión de las preguntas de investigación

P2.1.- ¿Cuáles son los diferentes escenarios de relación de ORBEA S.Coop. con los clientes/visitantes virtuales? Los diferentes escenarios de relación del caso de estudio, son los que se detallan a continuación.

En primer lugar, el **escenario 1** denominado *web*. En este escenario, el visitante virtual puede acceder al sitio web de ORBEA S.Coop. (www.orbea.com) y acceder a la web corporativa (página home, página contacto, etc.) de la empresa. En el caso de que disponga de una cuenta de usuario, puede realizar diferentes gestiones como realizar el pedido a través del B2C, activar la garantía de una bicicleta, etc. El visitante que no disponga de una cuenta de usuario, en cambio, puede navegar sin problemas por el sitio web pudiendo llegar incluso hasta el catálogo online del B2C, pero no podrá realizar el pedido. El visitante virtual que acceda a este escenario, tras la aparición de una página de configuración de idioma-país, va a poder acceder a los otros escenarios con facilidad. Es decir, el escenario Web tal y como se puede observar en la primera figura, tiene que como propósito que el visitante virtual acceda al catálogo online del B2C lo antes posible. Por ello, se puede considerar también que el catálogo online forma parte de la web corporativa.

En segundo lugar, está el **escenario 2**, llamado **tienda online (B2C)**. En este escenario, el visitante virtual puede acceder a la tienda online de que dispone ORBEA S.Coop. desde su página web. De esta manera, el acceso al producto ya no queda en una visita a una tienda o *dealer* sino que se globaliza el acceso. Se permite que a nivel mundial, cualquier persona pueda acceder al catálogo e interesarse por un producto y/o realizar un pedido.

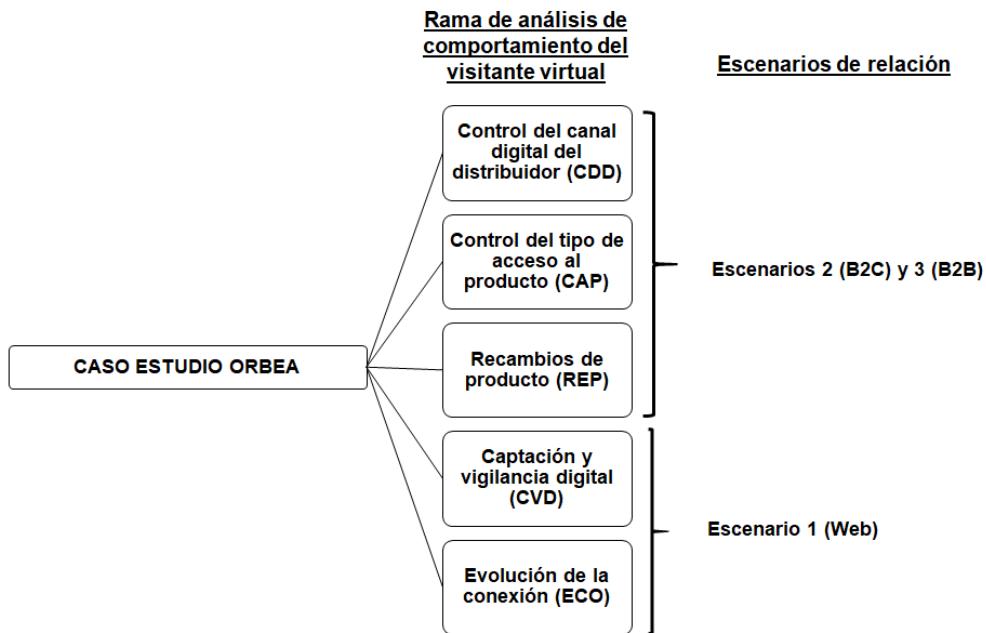
En tercer lugar, está el **escenario 3** denominado **plataforma B2B**. En este escenario, las tiendas o *dealers* tienen un rol diferente debido a la irrupción de la plataforma B2B, llamada KIDE. Dicha plataforma permite a las tiendas convertirse en visitante virtual, recordando que siguen siendo distribuidores de ORBEA S. Coop. y que las propias tiendas puedan acceder a toda la información de disponibilidad online y también puedan recibir los productos comprados igualmente

Capítulo 6. Conclusiones, limitaciones y líneas futuras

de manera online. Así, dejan de ser un mero distribuidor entre un usuario final y ORBEA S.Coop.

P2.2.- ¿Cuáles son las ramas de análisis de comportamiento del visitante virtual asociadas a cada escenario de relación? las ramas de análisis de comportamiento del visitante virtual asociadas a cada escenario de relación, explicados en la respuesta a la pregunta anterior, son las que se muestran en siguiente esquema (Figura 6.6).

Figura 6.6: Ramas de análisis de comportamiento del visitante virtual por escenario de relación. (Recapitulación)



Fuente: Elaboración propia

P2.3.- ¿Qué conclusiones se pueden obtener de la observación del comportamiento de los visitantes virtuales de forma aislada y comparativamente con las ventas realizadas? Las conclusiones se detallan por cada una de las ramas de análisis de comportamiento del visitante virtual:

En primer lugar, respecto a la Rama CDD, los visitantes virtuales tienen un interés mayoritario en bicicletas, concretamente de montaña y a los modelos ALMA, OIZ

6.2 Revisión de las preguntas de investigación

y RALLON. Además, el acceso que tienen los visitantes virtuales al escenario Plataforma B2B (KIDE) muestra, a nivel de accesos, lo siguiente. Por un lado, que los diferentes menús, salvo el de pedidos, tienen muy poco interés y por otro lado, que hay mucho acceso solamente al *login* y también a la raíz del propio escenario Plataforma B2B (KIDE). Siendo esto una oportunidad para que se fomenten entre los visitantes virtuales el resto del accesos a los menús del propio escenario Plataforma B2B (KIDE), que si son interesantes. De hecho esto último, está relacionado con una observación que se detectó en las reuniones, es decir, dentro de el escenario Plataforma B2B (KIDE) aunque la estructura de las direcciones de las páginas web sean amigables, existen un conjunto de movimientos, por parte de los visitantes virtuales, que facilitan romper dicha estructura. Por ejemplo, en la URL la pérdida de la palabra ”/kide/”, por lo cual, la parte del *Web Analytics* no puede registrar el movimiento en la navegación web, correctamente. Por ejemplo, en la dirección a las páginas web con el acceso al catálogo de productos dentro de una sesión del B2B, KIDE. Por consiguiente, esto requiere que el caso de estudio realice una mejora en el diseño de la navegación de su sitio web, debido a lo explicado anteriormente y que conlleva que los visitantes virtuales hagan otra ruta, la cual no está bien. Por otra parte, respecto al acceso a la activación de la garantía, Febrero de 2018 es un hito, ya que el número de registros de garantías comenzó a aumentar de manera considerable, alcanzando un crecimiento de más de un 100 % a los tres meses desde Febrero. Y respecto a los países con mayor actividad a nivel de acceso, copan las primeras posiciones: España, Francia, Reino Unido, USA e Italia.

En segundo lugar, en la Rama CVD, las organizaciones visitantes virtuales son en su mayoría visitantes potenciales. Además, la mayoría de la actividad de estos visitantes virtuales dónde su ISP es una empresa se recoge en Europa Central, USA y América Central.

Siguiendo con las ramas relacionadas con el acceso al producto, las Ramas REP y CAP, se observa en la parte de la ramas asociadas al producto del caso de estudio qué los visitantes virtuales en el caso de REP, tienen un interés principal por los recambios, aunque de 2017 a 2018 la actividad haya descendido. De hecho, en el

Capítulo 6. Conclusiones, limitaciones y líneas futuras

acceso a los recambios, los *paths* con mayor actividad son los asociados a las siguientes áreas de recambios: *equipaments* y *gear*. Y respecto a los países: España, USA, Reino Unido, Francia y Alemania copan las primeras posiciones en accesos en la rama REP. En cambio, en el caso de la rama CAP, los visitantes virtuales de tipo "Novel" son mayoritarios (y teniendo en cuenta toda la evolución siguen aumentando) en detrimento de los de tipo "Experto". En este sentido, en los meses de Julio de 2018 a Septiembre de 2018, es cuando los visitantes virtuales de tipo "Novel" han alcanzado una mayor actividad, respecto a toda la evolución temporal. Respecto a los *paths* asociados a los tipos de "Novel" y "Experto", en el primero, tienen mayor actividad los *paths* relativos a bicicletas de MTB, OIZ y ALMA y respecto al segundo, en cambio tienen mayor actividad los relativos a bicicletas de MTB y KIDS.

Asimismo, en la Rama ECO, predomina a nivel de IT la utilización del navegador web *Google Chrome* y *Safari* de Apple. Además, es interesante indicar que el acceso desde dispositivo móvil está en plena pugna con el acceso a escritorio, siendo el primero, el que va en cabeza. Asimismo, cabe destacar que en Julio de 2018 es dónde más ventaja ha obtenido. A nivel de Sistema Operativo, *Android* e *iOS* son también los que más se utilizan. Respecto a la categorización de *paths*, es el B2C el dominante con una diferencia considerable con el B2B y el "resto". A destacar, en los *paths* del B2C, el acceso a bicicletas de MTB como actividad mayoritaria desde España y Francia.

Finalmente, se observa comparativamente respecto a las ventas realizadas en ORBEA S.Coop, por un lado, qué los artículos más representativos son a primer nivel: BIKES y CLOTHING, a segundo nivel: MTB, KIDS y ROAD y a tercer nivel: MX, ALMA Y AVANT. Resaltar además que si se cruzan, por los artículos indicados en el interés online de la rama CDD, se concluye que a primer nivel los artículos de los pedidos y el interés online casi coinciden plenamente y que a segundo y tercer nivel hay algunas diferencias. Como por ejemplo, a tercer nivel ALMA se mantiene en ambos lados pero el resto no, pudiendo haber por lo tanto discrepancia entre los visitantes virtuales y los clientes reales. Por otro lado, respecto a los países, en ambos lados España y Francia mantienen su mayoría, pero

6.2 Revisión de las preguntas de investigación

resulta significativo que el Reino Unido es el tercero en importe de pedidos y en cambio, en el interés online está en el quinto lugar, ocurriendo lo contrario en el caso de USA. Además, los visitantes virtuales de Italia, tienen un interés mayor que la propia compra en sí. Es decir, los visitantes virtuales de Italia realizan más visitas al sitio web de ORBEA S.Coop. en comparación con los pedidos que se realizan desde la propia Italia. En este sentido, ocurre al revés con Alemania.

P3.1.- ¿Cuáles de los valores de las características de las tecnologías de la información (IT) se tendrán en consideración en los escenarios de relación? Las características IT del comportamiento web de los visitantes virtuales, que serán las referentes tanto a la hora de acceder a la plataforma B2B, como al catálogo (B2C), o al resto, se exponen a continuación, todo se basa en los resultados obtenidos a través del análisis *Web Mining*, que van a incluir las previsiones a seis periodos vista en base a los modelos de predicción generados.

En primer lugar, sobre las características IT para el escenario B2B. los análisis realizados para predecir los accesos a las características IT, van a mostrar la importancia que tienen sobre el caso de estudio.

- a) Dispositivos:** es necesario indicar que a nivel de dispositivo, el acceso desde dispositivos de escritorio o *Desktop* ampliará su dominio. En cambio, los accesos desde dispositivo móvil o *Mobile* y el acceso vía *Tablet* seguirán con su pugna.
- b) Navegadores web:** a nivel de navegador web, *Chrome* incrementará su dominio y *Firefox* mantendrá la segunda plaza, siguiéndole de cerca *Edge* que le arrebatará la posición a *Safari*.
- c) Sistemas Operativos:** por último, respecto a los SO, hay que decir que a nivel de escritorio *Windows* aumentará su ventaja respecto a *Macintosh* (el nombre que le da *Google Analytics* al SO de *Apple* para versiones de escritorio, comúnmente conocido como *Mac OS*) y *Chrome OS*. En cambio, sobre los dispositivos móviles y tablets *Android*, le superará a *iOS* mientras que *Windows Phone* sufrirá un descenso.

Capítulo 6. Conclusiones, limitaciones y líneas futuras

En segundo lugar, siguiendo con las características IT para el escenario B2C, a continuación se muestran sus respectivos análisis de predicción de las características IT, disponiendo también de las previsiones.

- a) Dispositivos:** A nivel de dispositivo, el tipo de dispositivo *Mobile* encabezará la pugna con el tipo de dispositivo *Desktop*, dicha pugna seguirá en un futuro pero con cierta distancia del primero. El acceso vía dispositivo *Tablet* aumentará su presencia pero muy lejos (alrededor de unos 200.000 accesos menos) de los dos primeros.
- b) Navegadores web:** A nivel de navegador web, *Chrome* aumentará su ventaja sobre *Safari* y *Firefox* respectivamente, que mantendrán las siguientes posiciones.
- c) Sistemas Operativos:** A nivel de SO, en los de tipo escritorio, *Windows* se mantendrá en cabeza pero *Macintosh* sufrirá una fuerte subida. En lo que respecta a los SO de móviles y tablets, *Android* aumentará su ventaja sobre *iOS*.

Por último, están las características IT para categoría "resto". Considerando este último, cómo todos los accesos a direcciones de páginas web, que no se contemplan en los escenarios del catálogo ni del B2B. A continuación, se va a mostrar de manera gráfica los resultados de los análisis realizados sobre las características IT.

- a) Dispositivos:** A nivel de dispositivo, el tipo de dispositivo *Mobile* ganará la pugna con el tipo de dispositivo *Desktop* y el acceso vía dispositivo *Tablet* ampliará algo su presencia pero muy lejos (alrededor de unos 100.000 accesos menos) de los que van en cabeza.
- b) Navegadores web:** A nivel de navegador web, *Chrome* incrementará su ventaja sobre *Safari* y *Firefox*, que seguirán en las siguientes dos posiciones.
- c) Sistemas Operativos:** A nivel de SO, en la versión de escritorio, *Windows* ampliará ligeramente su ventaja respecto a *Macintosh*. Sin embargo, en las versiones de móvil y tablets, *Android* aumentará su ventaja sobre *iOS*, el cual subirá

6.2 Revisión de las preguntas de investigación

ligeramente. Tras el análisis surgido de estas relaciones se ha identificado el factor de plataforma de la rama ECO o Plataforma ECO.

P3.2.- ¿Cuáles son las relaciones, tanto entre los tipos, como en el acceso a los productos visitados, en la web y los vendidos? las relaciones, tanto entre los tipos, como en el acceso a los productos visitados, en la web y los vendidos se detallan a continuación:

En primer lugar, respecto al interés online de la Rama CDD, el análisis realizado sobre la parte de CDD, muestra que por un lado, a nivel de interés online la categoría BIKES ampliará su liderazgo sobre los productos de nivel 1. En cambio, en el grupo perseguidor la categoría E-BIKES perderá fuerza respecto a las categorías PARTS y GEAR que pertenecen a la parte de recambios. Además, a nivel 2, los accesos a las bicicletas de la categoría ROAD alcanzará al acceso a las bicicletas de la categoría MTB y entrarán en una pugna mayor. La categoría PARTS, en cambio, quedará en un tercer lugar pero ampliará su presencia y por el contrario, la categoría KIDS perderá protagonismo, respecto a los anteriores. Asimismo, a nivel 3, el acceso a la categoría PARTS ampliará su ventaja y se pondrá en cabeza. El modelo ORCA se mantendrá en segundo lugar y el acceso al modelo de bicicleta TERRA superará al acceso al modelo de bicicleta OIZ. Sin embargo, el acceso al modelo de bicicleta ALMA quedará relegado a una quinta posición. Ya que la predicción prevee que no va a tener un aumento considerable en accesos por parte los visitantes virtuales. Resulta en cierta medida curioso el desplome que sufrirá el acceso al modelo de bicicleta GAIN.

En segundo lugar, en el Acceso al escenario Plataforma B2B (KIDE) de la Rama CDD, los accesos a otros menús o apartados irá creciendo más todavía en comparación con las secciones de pedidos, pedidos checking y content.

En tercer lugar, en el acceso a la garantía de la Rama CDD, habrá un punto de inflexión entre el acceso al formulario previo y el registro, siendo el primero quién tomará amplia ventaja respecto al segundo. Una vez visto el análisis surgido de

Capítulo 6. Conclusiones, limitaciones y líneas futuras

estas relaciones se ha identificado dentro de la parte del interés *online* el factor del equilibrio de la rama CDD o equilibrio CDD.

Por otra parte, en la Rama CAP, la previsión ha mostrado que aumentará la distancia entre el visitante virtual de tipo "novel" y el de tipo "experto", en lo referente al tipo de acceso al producto. A pesar de que este último también aumentará en el número de accesos. Aunque, en la serie histórica, se mantiene la ventaja del visitante virtual de tipo "novel" sobre el de tipo "experto". Tras el análisis surgido de estas relaciones, se ha identificado el factor de Madurez de la rama CAP o Madurez CAP.

Siguiendo con las ramas, en la REP, sobre los accesos a los recambios que realizarán los visitantes virtuales, es el propio acceso al recambio, el que mantendrá su ventaja respecto al acceso asociado al producto y al acceso asociado al modelo, respectivamente. No obstante, hay que indicar que el acceso al producto tomará una ligera ventaja respecto a los accesos asociado al modelo. Tras el análisis surgido de estas relaciones se ha identificado el factor de Tendencia de la rama REP o Tendencia REP.

Por último, se ha identificado el factor de proporción CVD relativo a la rama CVD, que debido a las limitaciones para su evolución en el proyecto actual, el desarrollo de la propia rama se ha limitado solamente al análisis heurístico presentado.

P4.1.- ¿Cuáles son las pautas para construir una estrategia analítica? Las pautas, por cada una de las ramas de análisis de comportamiento del visitante virtual, para construir una estrategia analítica son las siguientes:

En primer lugar, respecto a la Rama CDD (Control del Canal Digital del Distribuidor) se definen las siguientes premisas:

1. Interés online: acercar lo que querrán los visitantes virtuales con lo que ocurre hoy en día. Es decir, respecto al nivel 1 los visitantes virtuales seguirán

6.2 Revisión de las preguntas de investigación

queriendo comprar de la categoría BIKES. En cambio, a nivel 2, será el tipo ROAD el producto que los visitantes virtuales querrán comprar en dura pugna con el tipo MTB. A día de hoy MTB es el producto de nivel 2 más vendido y más visitado por los visitantes virtuales. Cabe resaltar, que la categoría KIDS que ahora es el segundo que más vende en un futuro perderá fuerza. Finalmente a nivel 3, el modelo ALMA y el modelo OIZ quedarán relegados en un futuro. El modelo ORCA se pondrá en un segundo lugar y a continuación el modelo TERRA.

2. Acceso al escenario Plataforma B2B (KIDE): potenciar los diferentes menús de la plataforma B2B KIDE. A día de hoy y también por lo que indica el análisis de predicción, la mayoría de los accesos no son a los menús de pedidos y pedidos checking sino que se orientan hacia otros sitios de la plataforma. Por lo tanto, se recomienda al caso de estudio qué oriente a los visitantes virtuales, a dichos menús a través de acciones vía desarrollo del sitio web.
3. Acceso al escenario Plataforma B2B (KIDE): no perder la referencia a la partes de la estructura de las direcciones de las páginas web, que se va generando a medida que los visitantes virtuales van navegando. Por ejemplo: Pérdida ”/kide/” en el catálogo. Rediseñar la arquitectura de navegación del sitio web.
4. Acceso a Garantía: establecer el registro de garantía de otra manera más visible y accesible, a través del desarrollo del sitio web, para el visitante virtual. Aunque a día de hoy, el acceso al registro ha aumentado, según la predicción va a bajar a futuro.

En segundo lugar, en la Rama CAP (Control de tipo de Acceso al Producto), se exponen las siguiente premisa:

5. Fomentar que los visitantes virtuales de acceso de tipo ”Experto” entren más al sitio web. Tanto hoy en día, como lo que indica la predicción realizada, el tipo de acceso ”Novel” domina y seguirá dominando el tipo de acceso al producto.

En tercer lugar, respecto a la Rama REP (Recambios de Productos) se muestra las siguiente premisa:

Capítulo 6. Conclusiones, limitaciones y líneas futuras

6. Facilitar que los visitantes virtuales que accedan a los modelos y productos del apartado de equipamientos. A día de hoy y por lo que indica la predicción el acceso a los recambios es y seguirá siendo el acceso mayoritario y con gran diferencia respecto a los otros dos.

Finalmente, en la Rama ECO (Evolución de la Conexión), las pautas son las siguientes:

7. Para el escenario Plataforma B2B (KIDE): fomentar que se pueda acceder por tecnología *Apple* (*iOS*, *Safari*, etc.) y vía APPs. Ya que a futuro, en base las predicciones realizadas, se prevé que el SO *Windows*, el dispositivo *Desktop* y el navegador web *Chrome* lideren dichas características IT respecto a los accesos que realizan los visitantes virtuales.
8. Para B2C: dedicar recursos a preparar el acceso al B2C para plataforma con SO *Android*, tipo de dispositivo principal *Mobile* y también para ordenador o *Desktop* y navegador web *Chrome* y *Safari*.

6.3 Conclusiones

En este apartado, se van a exponer las conclusiones del proyecto, a través de la explicación de la conclusión global, sobre la metodología aplicada y finalmente la conclusiones sobre los resultados.

6.3.1 Conclusión global

Como conclusión general y que tiene efecto sobre todas las preguntas de investigación es la elaboración de un modelo de comportamiento de visitantes virtuales para los proveedores de primer nivel del sector de distribución, habiendo cubierto de esta manera el *gap* identificado en la revisión de la literatura.

La relevancia de este proyecto de investigación afecta a dos ámbitos esenciales, como son el científico-académico y el empresarial. En el caso del primero, la elaboración de este proyecto de investigación ha permitido disponer de un modelo no

existente anteriormente para los proveedores de primer nivel del sector de la distribución. De esta manera, se ha cubierto el *gap*, sobre el comportamiento del visitante virtual en los proveedores de la cadena de suministro del sector de la distribución. En este sentido, hay que recalcar las evoluciones o comparativas realizadas durante el estado del arte (Tablas 2.1, 2.2, 2.3, 2.5, etc.). Además, este modelo ha permitido profundizar en la comprensión del comportamiento del visitante virtual generando un conjunto de elementos, factores y las relaciones entre ellos, que se muestran en el esquema del modelo. Todo ello, supone un punto de partida que futuros investigadores podrían aprovechar para generar más conocimiento y evolucionar el lado más científico del modelo.

En el caso del segundo, el desarrollo de esta tesis doctoral ha generado un modelo que gracias a la estrategia analítica que conlleva, los agentes del sector disponen de un conjunto de recomendaciones en base al comportamiento de sobre los visitantes virtuales, tanto para valorar inversiones en tecnología como para estrategias de marketing a poder llevar a cabo. Asimismo, el conocimiento que aportan los factores del modelo, les podrá servir a los agentes del sector para mejorar la utilización de los sistemas de comercio electrónico y poder influir a su vez en sistemas de gestión de las relaciones con los clientes.

6.3.2 Conclusión sobre la metodología aplicada

Tras haber elaborado este proyecto, es necesario indicar que la aplicación del enfoque cualitativo junto con el tipo de estudios de caso único junto con el diseño de la investigación realizado, ha permitido conocer el comportamiento del visitante virtual de los proveedores de primer nivel en el sector de la distribución y poder generar el modelo resultante, con la garantía de haber seguido el conjunto de pasos correcto.

6.3.3 Conclusiones sobre los resultados

Disclaimer: el contenido de este apartado no está disponible por cuestiones de derechos intelectuales/industriales.

Capítulo 6. Conclusiones, limitaciones y líneas futuras

6.4 Aportaciones

La relevancia de este proyecto de investigación afecta a dos ámbitos esenciales, como son el científico-académico y el empresarial. De esta manera, se ha cubierto el *gap*, sobre el comportamiento del visitante virtual en los proveedores de la cadena de suministro del sector de la distribución.

Para el ámbito científico-académico, este modelo ha permitido profundizar en la comprensión del comportamiento del visitante virtual generando un conjunto de elementos, factores y las relaciones entre ellos.

Para el ámbito empresarial, es importante recalcar que este proyecto tiene como uno de los propósitos principales crear una nueva línea de producto que redundará en una diversificación de la oferta actual. Incluyendo la mejora del ERP IZARO (producto de i68), ya que este proyecto va a ayudar a un crecimiento, desarrollando un nuevo producto que complementa y refuerza el impacto del nivel extendido del propio ERP. Además, también va a impactar en otros proyectos importantes de la unidad de I+D+i, al poder aplicar el conocimiento descubierto sobre los visitantes virtuales en el comercio electrónico. Sin olvidar que, el registro de propiedad intelectual del *software* realizado, así como su transferencia para explotación empresarial a la empresa del grupo internacional Zucchetti: Informática 68. Permitiendo así, su aplicabilidad en diferentes empresas como la del caso de estudio ORBEA S.Coop.

Una vez vistas las aportaciones, a continuación se van a describir las limitaciones y líneas futuras identificadas, ya que un indicador muy importante de la calidad de una investigación es su capacidad potencial para otros investigadores, así como, para apoyar futuras líneas de investigación, sin olvidar los inconvenientes que se han tenido en este proyecto.

6.5 Limitaciones y líneas futuras

En este apartado, se indican las limitaciones identificadas en el proyecto de investigación así como las líneas futuras que se proponen para futuras investigaciones.

6.5.1 Limitaciones

La limitación principal ha ocurrido en la parte del origen de datos de *Web Analytics*. Es decir, tras realizar la revisión de la literatura y en la propia metodología se estableció que *Google Analytics* y *Piwik* fueran los *trackers* que formasen la parte del *Web Analytics*. Sin embargo, tras un inicio prometedor pudiendo combinar ambas herramientas se nos comunicó por parte de la empresa que proporciona el servicio de *Piwik Cloud* que no querían seguir con el acuerdo alcanzado previamente, de suministrarnos una tarifa personalizada debido a que les justificamos que estábamos en un proyecto de investigación, y que si no pasábamos a una tarifa de producción debido al volumen de datos que se estaba generando, nos cortarían el servicio. Por lo cual, tras intentar reconducir la situación y ante la inviabilidad de encontrar una solución al problema se decidió desde la dirección del proyecto seguir solamente con *Google Analytics*, descartando otro *tracker Open Source* ya que cualquiera de ellos nos podría poner en la misma tesitura y además no todos tenían la arquitectura *cloud*.

6.5.2 Líneas futuras

En primer lugar, y más bien como mejora, se propone que se pueda aplicar la llamada generalización analítica, la cual permite que segmentos de estos resultados consigan emplearse en demás contextos (Williams et al., 2005). Debido a que el proyecto se ha basado en un estudio de caso. En este sentido, al referirnos a segmentos se quiere indicar que debido a que no hay dos contextos equivalentes, la transferencia nunca será completa.

En segundo lugar, se plantea refutar el estudio a distintos contextos, en los cuales, con otros casos se pueda comprobar si las respuestas a las cuestiones de investigación trazadas brindan puntos de coincidencia y los resultados son contrastables con

Capítulo 6. Conclusiones, limitaciones y líneas futuras

los conseguidos en nuestro proyecto de investigación. Permitiendo así, una mayor generalización analítica, ya que partiendo desde un caso único no tiene que significar comenzar desde una abstracción u algo utópico, sino que se puede considerar cómo un paradigma de la vida real (Giménez, 2012). Los contextos que se proponen son otros sectores como el de automoción, salud, etc. y otras áreas dentro del sector de distribución como por ejemplo el orientado a proveedores de otros ámbitos como alimentación, fabricación, etc.

En tercer lugar, este proyecto, al ser aplicado sobre un sitio web con sistemas de comercio electrónico, aporta conocimiento sobre el comportamiento de los visitantes o usuarios web que pueden ser de utilidad para otros proyectos importantes de la unidad de I+D+i de i68, relacionada con la gestión de las relaciones con el cliente, como por ejemplo el proyecto *CRM Alerts*, etc.

Aparte, a nivel más técnico del proyecto, se plantea que en futuros trabajos se pueda desarrollar la rama CVD, de análisis de comportamiento de visitante virtual, de CVD de manera más natural, con casos de estudio que dispongan de un sistema CRM, que contenga tipologías de clientes. Así, se podrán realizar segmentaciones de clientes en base a su comportamiento virtual de forma más precisa y exhaustiva, mostrando así la fertilidad de los resultados, ya que pueden prolongarse provocando en este caso el desarrollo que esta rama CVD no ha podido alcanzar en este proyecto de investigación.

Por último, se deja la puerta abierta a poder elaborar diversos proyectos de investigación de enfoque cuantitativo que trabajan a partir del modelo resultante generado en este proyecto.

APÉNDICE

A

Visualizaciones del Análisis Web Mining

Disclaimer: el contenido de este anexo no está disponible por cuestiones de derechos intelectuales/industriales.

Capítulo A. Visualizaciones del Análisis Web Mining

- AECOC. (2009a). *Pérdida desconocida*. Retrieved from <http://www.aecoc.es/?id=139&plantilla=11&target=Men%FA%3A%C1reas+de+actuaci%F3n> (Último acceso: 13/12/2015) 19, 20
- AECOC. (2009b). *Recomendación de prevención de pérdida desconocida* (Tech. Rep.). Asociación de fabricantes y distribuidores. 19, 21
- Agarwal, B. B., Khan, M., & Dhall, S. (2010, dec). Web Mining: Information and pattern discovery on the world wide web. In IJSTM (Ed.), *International journal of science, technology & management*; 54, 55
- Aggarwal, C. C. (2014). An introduction to frequent pattern mining. *Frequent Pattern Mining*, 1–17. Retrieved from http://dx.doi.org/10.1007/978-3-319-07821-2_1 doi: 10.1007/978-3-319-07821-2_1 54, 55
- Anderson, E., & Jap, S. D. (2005). The dark side of close relationships. *MIT Sloan Management Review*, 46(3), 75–82. 28, 31
- Association, W. A. (2008). *Web analytics definitions*. Retrieved from https://www.digitalanalyticsassociation.org/Files/PDF_standards/WebAnalyticsDefinitions.pdf (Último acceso: 02.10.2018) 41
- Azevedo, A. I. R. L. (2008). Kdd, semma and crisp-dm: a parallel overview. 52, 53
- Aziza, B. (2018). *Gartner magic quadrant for business intelligence (bi) 2018: The good, the bad, the ugly...* Retrieved from <https://www.atscale.com/blog/gartner-magic-quadrant-for-business-intelligence-bi-2018-the-good-the-bad-the-ugly> (Último acceso: 02.10.2018) 50
- Bal, J., Wilding, R. D., & Gundry, J. (1999). Virtual teaming in the agile supply chain. 29, 31
- Barbolla, J. O., Román, N. V., & Yagüe, M. J. (2013). El papel de las tics en la creación de beneficios en la cadena de valor de la distribución. In *Estratégias de distribución y comportamiento de compra multicanal: Tendencias y oportunidades para que fabricante y distribuidor rentabilicen sus decisiones de marketing* (pp. 109–132). 27, 31

-
- Belk, M., Papatheocharous, E., Germanakos, P., & Samaras, G. (2013). Modeling users on the world wide web based on cognitive factors, navigation behavior and clustering techniques. In *The journal of systems and software 86* (2013). doi: <http://dx.doi.org/10.1016/j.jss.2013.04.029> 1, 5
- Best Education, S. (2013). *The internet a decade later*. [bestedsites.com](http://www.bestedsites.com/the-internet-a-decade-later/). Retrieved from <http://www.bestedsites.com/the-internet-a-decade-later/> (Último acceso: 07.01.2016) 1, 5
- Bloom, J. Z. (2004, dec). Tourist market segmentation with linear and non-linear techniques. *Tourism Management*, 25(6), 723–733. 56
- Bonmatí, J. M. (2015, feb). *¿el año de la recuperación?* Retrieved from <http://sede.aecoc.es/fotos/marketing/web/actualidad%20economica.pdf> (Último acceso: 19/01/2016) 5, 7, 70
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32. 57
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Pacific Grove, California: Wadsworth. Inc. 57
- Burby, J., Brown, A., Committee, W. S., et al. (2007). Web analytics definitions. Washington DC: Web Analytics Association. 41
- Busto, I. (2012). *Desarrollo de un modelo de aplicación de las tics para la mejora de la integración de la cadena de suministro de los proveedores de componentes de primer nivel en el sector de automoción de la comunidad autónoma del país vasco* (Competitividad Empresarial y Desarrollo Económico). Universidad de Deusto. 16, 17, 25, 26
- Carazo, P. C. M. (2006). El método de estudio de caso: estrategia metodológica de la investigación científica. In (pp. 165–193). Ediciones Uninorte. 65, 77
- Castánder, I., Gerrikagoitia, J. K., & Alzua-Sorzabal, A. (2013, dic). Estudio comparativo de las herramientas de business intelligence: Empoderando el criterio de selección a las pymes. *tourGUNE Journal of Tourism and Human Mobility, Issue 1*, 45–56. Retrieved from https://www.researchgate.net/profile/Inigo_Castander/publication/259442931_Estudio_comparativo_de_las_herramientas_de_Business_Intelligence_Empoderando_el_criterio_de_selección_a_las_PYMES/links/00b4952b978220dc81000000.pdf 51
- Castánder, I., Serna, A., & Busto, I. (2017). Metodología para el comportamiento

Capítulo A. Visualizaciones del Análisis Web Mining

- de los visitantes virtuales en los proveedores del sector de la distribución. *Distribución y Consumo*, 4, 110-115. Retrieved from http://mercasa.es/files/multimedios/1509983957_Mercasa_distribucion_y_consumo_149_METODOLOGIA_100px.pdf 2, 7, 39, 40
- Chandrashekhar, A., & Schary, P. B. (1999). Toward the virtual supply chain: the convergence of it and organization. *The International Journal of Logistics Management*, 10(2), 27–40. 26
- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88–98. 49
- Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS quarterly*, 36(4), 1165–1188. 49
- Chen, L.-D., & Tan, J. (2004). Technology adaptation in e-commerce:: Key determinants of virtual stores acceptance. *European Management Journal*, 22(1), 74–86. 29, 31
- Chen, R.-C., Chiu, M.-L., Huang, Y.-L., & Chen, L.-T. (2004). Detecting credit card fraud by using questionnaire-responded transaction model based on support vector machines. In *Intelligent data engineering and automated learning–ideal 2004* (pp. 800–806). Springer. 58
- Chiva Gómez, R. (2001). El estudio de casos explicativos. una reflexión. *Revista de Economía y Empresa*, 15(41), 119–132. 66
- Christopher, M. (2000). The agile supply chain: competing in volatile markets. *Industrial marketing management*, 29(1), 37–44. 25
- Christopher, M., & Jüttner, U. (2000). Developing strategic partnerships in the supply chain: a practitioner perspective. *European Journal of Purchasing & Supply Management*, 6(2), 117–127. 32, 38
- Ciotti, G. (2015, may). *The top 10 best web analytic tools*. Retrieved from <http://www.sparringmind.com/best-web-analytics/> (Último acceso: 26/12/2015) 45
- Clark, D., Nicholas, D., & Jamali, H. R. (2014). Evaluating information seeking and use in the changing virtual world: the emerging role of google analytics. *Learned Publishing*, 27(3), 185–194. doi: 10.1087/20140304 42, 43
- Cloquell Lozano, A. (2015, abr). Usos sociales de internet entre los adolescentes españoles. *Revista sobre la infancia y la adolescencia*(8), 1. Re-

-
- trieved from <http://dx.doi.org/10.4995/reinad.2015.3649>
doi: 10.4995/reinad.2015.3649 2
- Constantinides, E. (2002). The 4s web-marketing mix model. *Electronic commerce research and applications*, 1(1), 57–76. 33, 34, 38
- Constantinides, E. (2004). Influencing the online consumer's behavior: the web experience. *Internet research*, 14(2), 111–126. 33, 34, 38
- Constantinides, E., & Geurts, P. (2005). The impact of web experience on virtual buying behaviour: an empirical study. *Journal of customer behaviour*, 4(3), 307–335. 33, 34, 38
- Constantinides, E., Schepers, L., & Vries, S. D. (2015). B2c social media value gap-model: a study of the dutch online retailing. *International journal of electronic marketing and retailing*, 6(3), 179–193. 36, 38
- Cullen, A. J., & Webster, M. (2007). A model of b2b e-commerce, based on connectivity and purpose. *International Journal of Operations & Production Management*, 27(2), 205–225. 36, 38
- Curry, J., & Curry, A. (2002). *CRM: cómo implementar y beneficiarse de la gestión de las relaciones con los clientes:(Customer Relationship Management)*. Gestión 2000.
- Davis, R. E., Elder, K., Howlett, D., & Bouzaglou, E. (1999). Relating storm and weather factors to dry slab avalanche activity at Alta, Utah, and Mammoth Mountain, California, using classification and regression trees. *Cold Regions Science and Technology*, 30(1), 79–89. 57
- De la Garza García, J., Morales Serrano, B., & González Cavazos, B. (2013). *Ánalisis estadístico multivariante*. McGraw Hill.
- Del Águila, A. R., Padilla, A., & Jiménez, J. A. (2000). Implicaciones estratégicas del comercio electrónico basado en internet: modelos de negocio y nuevos intermediarios. *Información Comercial Española, ICE: Revista de economía*(783), 63–78. 34, 38
- Duman, E., & Ozcelik, M. H. (2011). Detecting credit card fraud by genetic algorithm and scatter search. *Expert Systems with Applications*, 38(10), 13057–13063. 58
- Fasel, D., & Zumstein, D. (2009). *A fuzzy data warehouse approach for web analytics*. Springer. 48
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to know-

Capítulo A. Visualizaciones del Análisis Web Mining

- ledge discovery in databases. *AI magazine*, 17(3), 37. 42, 52
- Fernández de Pinedo, I. (2007). Construcción de una escala de actitudes tipo likert. *Centro de investigación y asistencia técnica. Recuperado de* [*http://www.siafa.com.ar/notas/nota164/escalera.htm*](http://www.siafa.com.ar/notas/nota164/escalera.htm).
- Fernando, D. (2014, jul). *10 web analytics tools: Analytical tools other than google analytics*. Retrieved from <http://blog.woorank.com/2014/07/10-web-analytics-tools-analytical-tools-other-than-google-analytics/> (Último acceso: 29/12/2015) 45
- Fernández Morales, I. (2014, dic). *La adopción y el uso de la tecnología por los mayores*. Responsable contenidos. Envejecimiento en red. Departamento de Población. CSIC. Retrieved from <https://envejecimientoenred.wordpress.com/2014/12/16/la-adopcion-y-el-uso-de-la-tecnologia-por-los-mayores/> 2
- Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2), 179–188. 56
- Franklin, C., & Ballan, M. (2005). Reliability and validity in qualitative research. In R. Grinnell & A. Unrau (Eds.), *Social work: Research and evaluation. quantitative and qualitative approaches* (7th ed., pp. 438–439). New York: Oxford University Press. 76
- Fu, C. Y. (2004). Combining loglinear model with classification and regression tree (CART): an application to birth data. *Computational Statistics & Data Analysis*, 45(4), 865–874. 57
- Garcia-Dastugue, S. J., & Lambert, D. M. (2003). Internet-enabled coordination in the supply chain. *Industrial marketing management*, 32(3), 251–263. 29, 31
- Gartner. (2015). *Gartner magic quadrant* (Methodologies). Author. Retrieved from http://www.gartner.com/technology/research/methodologies/research_mq.jsp 49
- Gereffi, G. (2001). Shifting governance structures in global commodity chains, with special reference to the internet. *American Behavioral Scientist*, 44(10), 1616–1637. 35, 38
- Ghauri, P. N., & Grønhaug, K. (2005). *Research methods in business studies: A practical guide*. Pearson Education. 66
- Ghenniwa, H., Dang, J., Huhns, M., & Shen, W. (2006). emarketplace model: an

-
- architecture for collaborative supply chain management and integration. In *Multiagent based supply chain management* (pp. 29–62). Springer. 30, 31
- Gillenson, M. L., Sherrell, D. L., et al. (2002). Enticing online consumers: an extended technology acceptance perspective. *Information & management*, 39(8), 705–719. 29, 31
- Giménez, G. (2012). El problema de la generalización en los estudios de caso. *Cultura y representaciones sociales*, 7(13), 40–62. 106
- Golicic, S. L., Davis, D. F., McCarthy, T. M., & Mentzer, J. T. (2002). The impact of e-commerce on supply chain relationships. *International Journal of Physical Distribution & Logistics Management*, 32(10), 851–871. 32, 38
- González, F. J. M., Mera, A. C., Lacoba, S. R., & Mogollón, L. R. (2006). Evaluación de la presencia en internet de los principales operadores logísticos en España. *Cuadernos de Estudios Empresariales*, 16, 99–114. 34, 38
- Google, D. (2016, jan). *Dimensions & metrics explorer*. Retrieved from <https://developers.google.com/analytics/devguides/reporting/core/dimsmets#mode=api&cats=ecommerce> (Último acceso: 12/01/2016) 47
- Gray, D. E. (2013). *Doing research in the real world*. Sage. 10, 62, 63
- Guerrero, M. (2015, nov). *Pérdida desconocida en comercios españoles registra índice más bajo desde 2010*. Retrieved from <http://noticiaslogisticaytransporte.com/logistica/17/11/2015/perdida-desconocida-en-comercios-espanoles-registra-indice-mas-bajo-desde-2010/62368.html> (Último acceso: 13/12/2015) 19, 21
- Gunasekaran, A., & Ngai, E. W. (2004). Information systems in supply chain integration and management. *European Journal of Operational Research*, 159(2), 269–295. 35, 38
- Gunasekaran, A., & Ngai, E. W. (2005). Build-to-order supply chain management: a literature review and framework for development. *Journal of operations management*, 23(5), 423–451. 28, 31
- Harrell, F. E., & Lee, K. L. (1985). A comparison of the discrimination of discriminant analysis and logistic regression under multivariate normality. *Biostatistics: Statistics in Biomedical, Public Health and Environmental Sciences?*, North-Holland, New York, United States, 333–343. 56

Capítulo A. Visualizaciones del Análisis Web Mining

- Hassini, E., Surti, C., & Searcy, C. (2012). A literature review and a case study of sustainable supply chains with a focus on metrics. *International Journal of Production Economics*, 140(1), 69–82. 28, 31
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *Unsupervised learning*. Springer. 56, 57
- Hausmann, V., Williams, S. P., & Schubert, P. (2012). Developing a framework for web analytics. *Master's Thesis, University of Koblenz-Landau*. 46, 47
- Hernández, B., Jiménez, J., & Martín, M. (2006). Análisis del efecto de la experiencia en tic en el uso del b2b. *Universidad de Zaragoza, available at: www.audem-virtual.com/articulos/122779009300.pdf (accessed 10 October 2010)*. (.) 36
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2006). Metodología de la investigación. *Editorial McGraw-Hill*,. 10, 62, 63
- Humphreys, P., McIvor, R., & Cadden, T. (2006). B2b commerce and its implications for the buyer-supplier interface. *Supply Chain Management: An International Journal*, 11(2), 131–139. 36, 38
- IBM, S. . D. (2012). Retail 2020: Reinventando la distribución minorista (una vez más). In *Un proyecto conjunto de ibm and y la new and york university and stern school and of business*. Retrieved from http://www-05.ibm.com/services/es/bcs/pdf/REW03013-ESES-01_HR.pdf 6, 15, 18, 19
- Indra, B. C. (2015). Retos estratégicos del fmccg y del retail.. Retrieved from http://businessconsulting.indracompany.com/sites/default/files/retos-estrategicos_AF_final.pdf 1, 6, 15, 17, 18, 19, 22, 23, 24
- INE. (2013). *El comercio electrónico y el uso de las nuevas tecnologías 2012* (Tech. Rep.). 2
- Iturbe-Ormaetxe, J. (2018). *Innovación de usuario en el sector de la bicicleta de montaña: el caso de la comunidad online de la orbea oiz en foromtb* (phdthesis). Mondragon Unibertsitatea. 6, 70
- James, M., Grosvenor, R., & Prickett, P. (2004). e-distribution: Internet-based management of a merchandiser supply chain. *Supply Chain Management: An International Journal*, 9(1), 7–15. 30, 31
- Jarvinen, J., & Karjaluoto, H. (2015, oct). The use of web analytics for di-

-
- gital marketing performance measurement. *Industrial Marketing Management*, 50, 117-27. Retrieved from <http://dx.doi.org/10.1016/j.indmarman.2015.04.009> doi: 10.1016/j.indmarman.2015.04.009 41, 46, 47
- Johnson, J. L. (1999). Strategic integration in industrial distribution channels: managing the interfirm relationship as a strategic asset. *Journal of the Academy of Marketing Science*, 27(1), 4–18. 32, 38
- Kaur, S., & Khanna, N. (2014, apr). Evaluation of web mining approach. In www.IJCSN.org (Ed.), *Ijcsn international journal of computer science and network* (Vol. 3). ISSN (Online) : 2277-5420. 53
- Kay Olsen, W., Warde, A., & Martens, L. (2000). Social differentiation and the market for eating out in the UK. *International Journal of Hospitality Management*, 19(2), 173–190. 56
- Kim, J.-C., Kim, D.-H., Kim, J.-J., Ye, J.-S., & Lee, H.-S. (2000). Segmenting the Korean housing market using multiple discriminant analysis. *Construction Management & Economics*, 18(1), 45–54. 56
- Kull, T., & Closs, D. (2008). The risk of second-tier supplier failures in serial supply chains: implications for order policies and distributor autonomy. *European Journal of Operational Research*, 186(3), 1158–1174. 28, 31
- Kurt, I., Ture, M., & Kurum, A. T. (2008). Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease. *Expert Systems with Applications*, 34(1), 366–374. 57
- Labrecque, L. I., Mathwick, C., Novak, T. P., Hofacker, C. F., et al. (2013). Consumer power: Evolution in the digital age. *Journal of Interactive Marketing*, 27(4), 257–269. 2, 6
- Lagus, K., Kaski, S., & Kohonen, T. (2004). Mining massive document collections by the WEBSOM method. (September 2004). 56
- Lai, K., & Cerpa, N. (2001). Support vs Confidence in Association Rule Algorithms. 58
- Lambert, D. M., & Cooper, M. C. (2000). Issues in supply chain management. *Industrial marketing management*, 29(1), 65–83. 32, 35, 38
- Lambert, D. M., & Schwieterman, M. A. (2012). Supplier relationship manage-

Capítulo A. Visualizaciones del Análisis Web Mining

- ment as a macro business process. *Supply Chain Management: An International Journal*, 17(3), 337–352. 35, 38
- Lancioni, R., Schau, H. J., & Smith, M. F. (2003). Internet impacts on supply chain management. *Industrial Marketing Management*, 32(3), 173–175. 29, 31
- Lancioni, R. A., Smith, M. F., & Oliva, T. A. (2000). The role of the internet in supply chain management. *Industrial Marketing Management*, 29(1), 45–56. 32, 38
- Lee, H. L., & Whang, S. (2004). E-business and supply chain integration. In *The practice of supply chain management: Where theory and application converge* (pp. 123–138). Springer. 29, 31, 32, 33, 38
- Liu, Q., & Zhang, A. M. (2013, sep). Application of Web Mining in Electronic Commerce. *Applied Mechanics and Materials*, 409-410, 1597–1603. Retrieved from <http://dx.doi.org/10.4028/www.scientific.net/AMM.409-410.1597> doi: 10.4028/www.scientific.net/amm.409-410.1597 55
- Liu, Y., Tao, L., Li, Y., & El-Ansary, A. I. (2007). The impact of a distributor's trust in a supplier and use of control mechanisms on relational value creation in marketing channels. *Journal of Business & Industrial Marketing*, 23(1), 12–22. 28, 31
- Lorenzo, C., Constantinides, E., & Gomez-Borja, M. A. (2009). Effects of web experience factors on virtual retail purchase preferences. 34, 38
- López Moraleda, R. (2014, abr). *La brecha digital, nueva causa de exclusión social de los mayores*. Retrieved from http://www.revista60ymas.es/60mas_01/informe/IM_084758?dDocName=IM_084758 2
- Manthou, V., Vlachopoulou, M., & Folinas, D. (2004). Virtual e-chain (vec) model for supply chain collaboration. *International Journal of Production Economics*, 87(3), 241–250. 25, 26
- Masterov, D. V., Mayer, U. F., & Tadelis, S. (2015). Canary in the e-commerce coal mine: Detecting and predicting poor experiences using buyer-to-seller messages. In *Proceedings of the sixteenth ACM conference on economics and computation* (pp. 81–93). 37, 38
- Maxwell, J. (1998). Handbook of applied social research method. In L. B. D. J. & Rog (Eds.), (p. 69-100). Thousand Oaks, CA, Sage. 77

-
- Mertens, D. M. (2014). *Research and evaluation in education and psychology: Integrating diversity with quantitative, qualitative, and mixed methods* (4th ed.). Sage Publications. 77
- Monge, E. C. (2010). El estudio de casos como metodología de investigación y su importancia en la dirección y administración de empresas. *Revista Nacional de administración*, 1(2), 31–54. 64, 65, 67, 76
- Moyano-Fuentes, J. (2010). La influencia de la integración de la cadena de suministro en la implantación de capacidades de comercio electrónico*/the influence of the supply chain integration on the level of implementation of e-business capabilities. *Universia Business Review*(28), 30. 35, 38
- Murtaza, M. B., Gupta, V., & Carroll, R. C. (2004). E-marketplaces and the future of supply chain management: opportunities and challenges. *Business Process Management Journal*, 10(3), 325–335. 33, 38
- Naresh Kumar Kar, A. A., H. R. Sharma. (2015). Knowledge retrieval from web server logs using web usage mining. In (Vol. Volume 4 Issue 3, p. 2319-7064). 42
- Negash, S., & Gray, P. (2008). *Business intelligence*. Springer. 49
- of Supply Chain Management Professionals (CSCMP), C. (2018). *Definitions of supply chain management*. Retrieved from https://cscmp.org/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms.aspx?hkey=60879588-f65f-4ab5-8c4b-6878815ef921
(Último acceso: 14.10.2018) 16
- ONTSI. (2013). *Estudio sobre Comercio Electrónico B2C 2012* (Tech. Rep.). Observatorio Nacional de las Telecomunicaciones y de la Sociedad de la Información. 2, 6
- ONTSI. (2015, Oct). *Estudio sobre comercio electrónico b2c 2014* (Tech. Rep. No. ISSN 2172-458X). ONTSI (Observatorio Nacional de las Telecomunicaciones y de los Sistemas de Información). 2, 6
- Pant, S., Sethi, R., & Bhandari, M. (2003). Making sense of the e-supply chain landscape: an implementation framework. *International Journal of Information Management*, 23(3), 201–221. 26
- Pettigrew, A. M. (1990). Longitudinal field research on change: Theory and practice. *Organization science*, 1(3), 267–292. 66

Capítulo A. Visualizaciones del Análisis Web Mining

- Piatetsky-Shapiro, G. (2007). Data mining and knowledge discovery 1996 to 2005: overcoming the hype and moving from “university” to “business” and “analytics”. *Data Mining and Knowledge Discovery*, 15(1), 99–105. 52, 53, 59
- Piwik, D. (2016). *Database schema*. Retrieved from <http://developer.piwik.org/guides/persistence-and-the-mysql-backend> (Último acceso: 12/01/2016) 47
- Porter, M. E., & Millar, V. E. (1985). *How information gives you competitive advantage*. Harvard Business Review, Reprint Service. Retrieved from <http://faculty.yu.edu.jo/iaad/Lists/Taught%20Courses/Attachments/5/Reading%205-How%20Information%20Gives%20You%20Comp-Fall2015.pdf> 17
- Rathod, M. D. B., Prajapati, M. R., & Khanna, D. S. (2015, apr). Emerging trends and of clustering and techniques in and web. In *Ijsart* (Vol. 1). 56
- Roche, I. C., Yagüe, M. J., Nogales, A. F., Suárez, M. G., Arévalo, A. R., Orta, O. M., ... Barbolla, J. O. (1999). *Los canales de distribución de productos de gran consumo: concentración y competencia*. Ediciones Pirámide. 32, 38
- Rodríguez, J. A. O. (2002). *Proceso de modernización de la cadena de suministro, una empresa tradicional preparandose para trabajar en ambientes de comercio electrónico-edición única* (Unpublished doctoral dissertation). 32, 38
- Rubio, N., & Yagüe, M. J. (2009). Alternative panel models to evaluate the store brand market share: Evidence from the spanish market. *European Journal of Marketing*, 43(1/2), 110–138. 27, 31
- Rubio Gil, A. (2010). Jóvenes en red, generación digital y cambio social.. 2
- Sakthipriya, C., Srinaganya, G., & Sathiaseelan, J. (2015). An analysis of recent trends and challenges in Web Usage Mining applications. 53, 54, 56
- Salcedo Barriga, L. (2012). Análisis del sector de la distribución: evaluación de la competitividad y situación actual. In *Universidad de comillas*. 7, 19, 22
- Sánchez, D., Vila, M. A., Cerdá, L., & Serrano, J.-M. (2009). Association rules applied to credit card fraud detection. *Expert Systems with Applications*, 36(2), 3630–3640. 58
- Sánchez, H., & Reyes, C. (2006). Metodología y diseños en la investigación científica.

-
- Serve, M., Yen, D. C., Wang, J.-C., & Lin, B. (2002). B2b-enhanced supply chain process: toward building virtual enterprises. *Business Process Management Journal*, 8(3), 245–253. 35, 38
- Sánchez Valle, M., & De Frutos Torres, B. (2012). Adolescentes en las redes sociales ¿marca el género la diferencia? In *Revista telos (cuadernos de comunicación e innovación)* - www.telos.es. 2
- Srivastava, J., Cooley, R., Deshpande, M., & Tan, P.-N. (2000). Web usage mining: Discovery and applications of usage patterns from web data. *ACM SIGKDD Explorations Newsletter*, 1(2), 12–23. doi: 10.1145/846183.846188 42, 46, 53
- Suh, E. H., Noh, K. C., & Suh, C. K. (1999). Customer list segmentation using the combined response model. *Expert Systems with Applications*, 17(2), 89–97. 56
- Sulbhewar, A. R., & Raut, P. A. D. (2015). Review paper and on web mining and its application. In I. . www.ijrise.org—editor@ijrise.org [120-124] (Ed.), *International journal of research in science & engineering* (Vol. 1). 54
- Sutton, R. I. (1997). Crossroads—the virtues of closet qualitative research. *Organization Science*, 8(1), 97–106. 76
- Tadelis, S. (2016). The economics of reputation and feedback systems in e-commerce marketplaces. *IEEE Internet Computing*, 20(1), 12–19. 37, 38
- Tafur Portilla, R. (1995). La tesis universitaria. *Editorial Mantaro. Lima*.
- Tamaddoni Jahromi, A., Stakhovych, S., & Ewing, M. (2014, oct). Managing B2B customer churn, retention and profitability. *Industrial Marketing Management*, 43(7), 1258–1268. Retrieved from <http://dx.doi.org/10.1016/j.indmarman.2014.06.016> doi: 10.1016/j.indmarman.2014.06.016 23
- Tang, C., & Tomlin, B. (2008). The power of flexibility for mitigating supply chain risks. *International Journal of Production Economics*, 116(1), 12–27. 28, 31
- Thomas, A., & Barton, R. (2007). Integrating local suppliers in a global supply network. *Journal of Manufacturing Technology Management*, 18(5), 490–512. 35
- Trappey, C. V., & Trappey, A. J. (1998). A chain store marketing information system: realizing internet-based enterprise integration and electronic commerce.

Bibliografía

- Industrial Management & Data Systems*, 98(5), 205–213. 32, 38
- Trevino, L. J., & Daniels, J. D. (1995). FDI theory and foreign direct investment in the United States: a comparison of investors and non-investors. *International Business Review*, 4(2), 177–194. 56
- Vapnik, V. (1998). *Statistical learning theory*. 1998. Wiley, New York. 58
- Vellingiri, J., Kaliraj, S., Satheeshkumar, S., & Parthiban, T. (2015, feb). A novel approach for user navigation pattern discovery and analysis for web usage mining. *Journal of Computer Science*, 11(2), 372-382. Retrieved from <http://dx.doi.org/10.3844/jcssp.2015.372.382> doi: 10.3844/jcssp.2015.372.382 46, 47
- Velásquez, J. D., & Donoso, L. (2010, jun). Aplicación de técnicas de web mining sobre los datos originados por usuarios de páginas web. In *Revista ingeniería de sistemas* (Vol. XIV). 53, 56
- Watson, H. J., & Wixom, B. H. (2007). The current state of business intelligence. *Computer*, 40(9), 96–99. 49
- Wehrens, R., & Buydens, L. M. C. (2007). Self-and super-organizing maps in R: the Kohonen package. *Journal of Statistical Software*, 21(5), 1–19. 56
- Williams, M., Tutty, L., & Grinnell, R. (2005). The qualitative research approach. In R. Grinnell & A. Unrau (Eds.), *Social work: Research and evaluation. quantitative and qualitative approaches* (7th ed., pp. 75–87). New York: Oxford University Press. 77, 105
- Wu, D., & Olson, D. L. (2008). Supply chain risk, simulation, and vendor selection. *International journal of production economics*, 114(2), 646–655. 28, 31
- Yagüe, M. J., & Rubio, N. (2006). Resultados económicos y relacionales del fabricante con la marca de distribuidor. *Distribución y consumo*(88), 72–85. 27, 31
- Yin, R. K. (1994). *Case study research: Design and methods*. Sage publications. 64, 65, 66, 77
- Zhang, K. Z., Zhao, S. J., Cheung, C. M., & Lee, M. K. (2014, nov). Examining the influence of online reviews on consumers decision-making: A heuristic–systematic model. *Decision Support Systems*, 67, 78–89. Retrieved from <http://dx.doi.org/10.1016/j.dss.2014.08.005> doi: 10.1016/j.dss.2014.08.005 5

Glosario

- AECOC** Asociación de Empresas del Gran Consumo 7
- AFIDAD** Asociación Española de Distribución Deportiva 71
- AMBE** Asociación de Marcas y Bicicletas de España 71
- ANGED** Asociación Nacional Grandes de Empresas de Distribución 71
- ANN** Artificial Neural Network. Es una Técnica de Data Mining. xii, 55
- API** Application Programming Interface 48
- AR** Association Rules. Es una Técnica de Data Mining. xii, 57
- B2B** Business to Business. Es el medio para realizar acciones comerciales a través de Internet entre diferentes organizaciones 6
- BI** Business Intelligence. Inteligencia de Negocio. Es un tipo de proceso que lleva a cabo un conjunto de etapas que van desde el acceso al origen del dato hasta su análisis o posterior consumo pasando por su extracción, tratamiento y estructuración. 32
- BRICS** Brasil, Rusia, India, China y Sudáfrica 6
- CART** Classification And Regression Tree xii, 57
- CRISP-DM** CRoss-Industry Standard Process for Data Mining 52
- CSCMP** Council of Supply Chain Management Professionals 16
- Customer Churn** Pérdida de clientes 23
- CV** Cross-Validation. Es una técnica de evaluación. 57
- DAA** Digital Analytics Association 41
- Dashboard** Panel o Cuadro de Mando. Se refiere a un documento en el que se reflejan, mediante una representación gráfica, las principales métricas o KPI que

Glosario

intervienen en la consecución de los objetivos de una estrategia. De ahí la importancia y su significado estratégico. 79

DM Data Mining. Minería de datos. Es el conjunto de técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas, clasificaciones, etc. que expliquen el comportamiento de los datos en un determinado contexto. 33

e-Supply Chain e-Supply Chain (Virtual Supply Chain) 25

ERP Enterprise Resource Planning. Es un sistema de información integrado de gestión empresarial. 99

FMI Fondo Monetario Internacional 15

GA Genetic Algorithms. Es una Técnica de Data Mining. xii, 58

GPL General Public License. 45

HTTP HyperText Transfer Protocol 42

Internet Red de redes o conjunto descentralizado de redes de comunicación interconectadas. 1

IR Interquartil Range. Rango intercuartil. 51

ISP Internet service provider. 92, 114

IT Information Technology. Tecnología de la Información 9, 69, 136

KDD Knowledge Discovery in Databases 42

LDA Linear Discriminant Analysis. Es una Técnica de Data Mining. xii, 56

LDF Linear Discriminant Function. Es una Técnica de Data Mining. 56

LR Logistic Regression. Es una Técnica de Data Mining. xii, 56

PIB Producto Interior Bruto 7

PYMES Pequeñas y Medianas Empresas 22

RF Random Forest. Es una Técnica de Data Mining. xii, 57

SCM Supply Chain Management. Gestión de la cadena de suministro. 16

SEMMA Sample, Explore, Modify, Model, Assess 52

- shrinkage** Pérdida desconocida 20
- SOM** Self-Organizing Mapping xii, 55
- SVM** Support Vector Machines. Es una Técnica de Data Mining. xii, 58
- TIC** Tecnologías de la Información y la Comunicación 5
- tradicionales** Perfil tradicional como segmento digital de la sociedad. 24
- URL** Uniform Resource Locator 99
- USA** Estados Unidos de América 6
- Web Analytics** El seguimiento de objetivos, recogida, medición, reporte y análisis de los datos cuantitativos de Internet para optimizar los sitios web y las iniciativas de marketing web 7
- Web Mining** Concepto que reúne todas las técnicas, métodos y algoritmos utilizados para extraer información y conocimiento a partir de datos procedentes de los sitios web. 53
- Weka** Software de Data Mining de tipo open-source 58
- WUM** Web Usage Mining 53

