



Universidad
Carlos III de Madrid

TRABAJO FIN DE GRADO

Curso 2015/2016

Optimización en Facebook de contenidos
especializados en el sector alimentario.

Creación de la empresa LULAF.SL

Autor: Luis Lacadena Pascual

Tutor Informática: Jorge Luis Morato Lara

Tutor ADE: Santiago Sánchez Cervera Serna

Doble Grado en Informática y ADE.

Madrid, Septiembre 2016

Resumen

Este proyecto tiene como objetivo crear una empresa que desarrolle una herramienta capaz de rastrear términos clave en la red social Facebook. Más concretamente, esta empresa ha desarrollado una herramienta para el mercado de la venta de productos alimenticios por Internet. La mayoría de las herramientas SEO no buscan información en las páginas de las redes sociales de un determinado campo. Esta herramienta en particular, busca los términos relevantes en las páginas más importantes del sector de la alimentación en internet. Se han analizado estas páginas y se han extraído los comentarios relevantes que deben ser analizados para devolver al cliente el resultado de la frecuencia de uso a lo largo de un periodo de tiempo. Esta herramienta puede servir a empresas y/o particulares a escoger los términos relevantes para sus páginas y la estacionalidad que éstos tienen mediante un servicio web que mostrará los resultados gráficamente.

Índice de ilustraciones

Ilustración 1. PAGERANK Fuente [1]	14
Ilustración 2. MOZRANK. Fuente [3]	16
Ilustración 3. Scrum iteración. Fuente [8]	20
Ilustración 4. Ejemplo de metodologías ágiles. Fuente [9]	21
Ilustración 5. Operador google en SI. Fuente propia	23
Ilustración 6. Categorías de consumo online. Fuente [17]	25
Ilustración 7. Compradores online franjas. Fuente [19]	26
Ilustración 8. API Facebook. Fuente [79].	42
Ilustración 9. Casos de uso. Elaboración propia.....	61
Ilustración 10. Diagrama de componentes	65
Ilustración 11. Capas de los sistemas. Fuente [66]	66
Ilustración 12. Diagrama de clases. Fuente: elaboración propia	78
Ilustración 13. Diseño de datos. Fuente: elaboración propia	80
Ilustración 14. Panel de control XAMPP. Fuente: elaboración propia	81
Ilustración 15. Demostración de URL. Fuente: elaboración propia	81
Ilustración 16. PHPmyAdmin. Fuente: elaboración propia	82
Ilustración 17. Subir datos a XAMPP. Fuente: elaboración propia	82
Ilustración 18. Ver fichero de carga. Fuente: elaboración propia.....	83
Ilustración 19. Selección de carga de tablas. Fuente: elaboración propia.....	83
Ilustración 20. Visión de términos clave. -Fuente propia	105
Ilustración 21. Fórmula ROE. Fuente [76]	110
Ilustración 22. Tasa de crecimiento del PIB. Fuente [33].....	112
Ilustración 23. Gráfica del paro. Fuente [34]	112
Ilustración 24. Previsión del paro. Fuente [35]	113
Ilustración 25. Piramide de población. Fuente [37]	113
Ilustración 26. Tabla de esperanza de vida Edad/Países. Fuente [38]	114
Ilustración 27. Porter. Fuente [43]	116
Ilustración 28. Precios de la competencia. Fuente Elaboración propia.	119
Ilustración 29. Activo fijo no corriente. Fuente propia	125
Ilustración 30. Evolución del PIB. Fuente [72].....	127
Ilustración 31. Evolución ventas online. Fuente [73]	128
Ilustración 32. Planificación tiempos. Fuente propia.....	133
Ilustración 33. Planificación gráfica. Fuente propia	134
Ilustración 34. Uso redes sociales. Fuente [78].....	135

Índice de tablas

Tabla 1 Resumen de tecnologías.....	22
Tabla 2- Carga posts inicial.....	44
Tabla 3- Carga diccionario palabras relevantes	44
Tabla 4- Carga de datos de palabras negativas.....	44
Tabla 5- Mostrar tabla de frecuencias	45
Tabla 6- Usuario debe introducir palabra deseada.....	45
Tabla 7- Mostrar palabras relevantes categoría Alimentación.....	46
Tabla 8- Mostrar palabras relevantes categoría Bebidas.....	46
Tabla 9- Mostrar palabras relevantes categoría Regiones.....	47
Tabla 10- Mostrar palabras relevantes categoría Términos especiales.....	47
Tabla 11- Mostrar palabras relevantes categoría Regiones.....	47
Tabla 12- Determinar la relevancia de un comentario (post).....	48
Tabla 13- Analizar frecuencia de post relevante.....	48
Tabla 14- Mostrar página de índice.....	49
Tabla 15- Usuario puede aconsejar páginas de Facebook a analizar.....	49
Tabla 16- Sistema operativo Windows.....	50
Tabla 17 - Procesador recomendado	50
Tabla 18- Memoria RAM recomendada	51
Tabla 19- El tiempo de respuesta.....	51
Tabla 20- Disponibilidad anual	52
Tabla 21- Claridad de diseño	52
Tabla 22- Portabilidad del sistema entre sistemas operativos	53
Tabla 23- Interacción entre los distintos subsistemas	53
Tabla 24- Interacción entre los distintos subsistemas	54
Tabla 25- Pantalla inicial.....	54
Tabla 26- Pantalla con gráfica de resultados	55
Tabla 27- Pantalla con términos de Bebidas	55
Tabla 28- Pantalla con términos de Alimentos	56
Tabla 29 - Pantalla con términos de Regiones de España.....	56
Tabla 30- Pantalla con términos de Bebidas	57
Tabla 31- Pantalla con términos de Comidas diarias	57
Tabla 32- Pantalla con términos de Comidas diarias	58
Tabla 33- Pantalla con términos de Comidas diarias	58
Tabla 34- Comunicación entre subsistemas.....	59
Tabla 35- Compatibilidad con otras aplicaciones.....	59
Tabla 36- C01.....	62
Tabla 37- C02.....	62
Tabla 38-C03.....	63
Tabla 39- C04.....	63
Tabla 40- C05.....	63
Tabla 41- Matriz de trazabilidad	64
Tabla 42- Guarda datos	69

Tabla 43- Simpleline	70
Tabla 44- Index	70
Tabla 45- PáginaAlimentación.....	71
Tabla 46- Ejemplo6(bebidas).....	71
Tabla 47-PáginaRegiones	72
Tabla 48- PáginaTérminosEspeciales	72
Tabla 49-PáginaTiposComida	73
Tabla 50-PresentaciónTabla	73
Tabla 51- ServletFormulario	74
Tabla 52- Comentario.....	75
Tabla 53- ListaComidas.....	75
Tabla 54- PalabrayFrecuencias.....	76
Tabla 55- PalabrasNegativas	76
Tabla 56- TextDirectoryToArff.....	76
Tabla 57- WriteFile	77
Tabla 58- PZN	84
Tabla 59- PF01	85
Tabla 60- PF02	85
Tabla 61-PF03	85
Tabla 62-PF04.....	86
Tabla 63-PF05.....	86
Tabla 64-PF06.....	86
Tabla 65-PF07	87
Tabla 66-PF08.....	87
Tabla 67-PF09	87
Tabla 68-PF10.....	88
Tabla 69-PF11	88
Tabla 70-PF12.....	89
Tabla 71-PA01.....	89
Tabla 72-PA02.....	89
Tabla 73 -PA03.....	90
Tabla 74-PA04.....	90
Tabla 75-PA05.....	90
Tabla 76-PA06.....	91
Tabla 77-PA07.....	91
Tabla 78-PA08.....	91
Tabla 79-PA09.....	92
Tabla 80-PA10.....	92
Tabla 81-PA11.....	93
Tabla 82-PA12.....	93
Tabla 83-PA13.....	93
Tabla 84-PA14.....	94
Tabla 85-PNF01	94
Tabla 86-PNF02	95
Tabla 87-PNF03	95
Tabla 88-PNF04	95
Tabla 89-PNF05	96
Tabla 90-PNF06	96

Tabla 91–PNF07	97
Tabla 92–PNF08	97
Tabla 93–PNF09	97
Tabla 94–PNF10	98
Tabla 95- Término y repeticiones.....	101
Tabla 96- Término y ponderación	104
Tabla 97- Creación de la empresa.....	106
Tabla 98- P&L	128
Tabla 99- Activos	129
Tabla 100- Pasivos.....	130
Tabla 101- Cash Flow.....	130
Tabla 102- Salidas de efectivo.....	131
Tabla 103- Punto Muerto	131
Tabla 104 -VAN.....	132
Tabla 105- Company description	142
Tabla 106. Cuenta de pérdidas y ganancias 16 al 18	155
Tabla 107. Entradas y salidas de efectivo, 16 al 18.....	156
Tabla 108. Datos de los gastos de la empresa	156
Tabla 109. Balance de cuentas.....	157
Tabla 110. Tabla de justificación del VAN	157
Tabla 111. Tabla de justificación de la TIR	158
Tabla 112. Tabla de balance de ene-16.....	158
Tabla 113. Justificación de estados de flujo de caja ene-16	158
Tabla 114. Justificación de los ingresos en 16.....	158
Tabla 115. Cálculo del punto muerto justificado.	159
Tabla 116. Justificación del ROE y del ROI.....	159
Tabla 117. Balance en 2016	159
Tabla 118. Estado de flujos de efectivo dic-16.....	160
Tabla 119. Balance de dic -17.....	160
Tabla 120. Estado de flujos de efectivo de di-17	160
Tabla 121. Balance de dic-18.....	161
Tabla 122. Estado de flujos de efectivo dic-18.....	161

Contenido

1.	Introducción	10
1.1	Planteamiento de la cuestión.....	10
1.2	Objetivos	11
1.3	Contenido de la memoria.....	12
2.	Estado del Arte	13
2.2	El algoritmo PAGERANK.	14
2.2	Otras herramientas para analizar la popularidad	15
2.3	Tecnologías necesarias para el desarrollo.....	16
2.4	Google Hacking:.....	22
2.5	El mercado de las ventas online de alimentos	25
2.6	Investigación sobre las mejores tiendas del año 2015.	27
2.7	Marco regulador.....	38
3	Análisis de la solución	38
3.2	Definición del sistema	39
3.3	Requisitos	42
3.4	Identificación de subsistemas de análisis.....	64
4	Diseño.....	65
4.1	Arquitectura del sistema	65
4.2	La interfaz de usuario	67
4.3	Identificación y especificación de clases	67
4.4	Diseño físico de datos.....	79
4.5	Migración de datos inicial.	80
5	Especificación técnica del plan de pruebas.....	83
5.1	Pruebas funcionales	84
5.2	Pruebas de aceptación.	89
5.3	Pruebas no funcionales	94
6	Micro lenguaje del segmento de la alimentación	98
6.1	Lista de términos.....	98
6.2	Los términos más significativos para el micro lenguaje.....	103
7	Taxonomía del micro lenguaje:.....	104
8	Descripción de la empresa.	106
8.1	Perfil de la empresa.....	106
8.2	Descripción del proceso productivo.....	107
9	Análisis de la situación interna de la empresa.	108
9.1	Capacidades personales.....	108

9.2	Capacidades técnicas	108
9.3	Capacidades comerciales	109
9.4	Capacidades de gestión	109
9.5	Capacidades financieras	109
10	Análisis de la situación externa de la empresa	110
10.1	Análisis PEST. Variables del entorno global	110
10.2	Análisis de las 5 fuerzas de Porter. Variables del entorno específico	115
11	Análisis DAFO	119
11.1	Oportunidades	119
11.2	Fortalezas	120
11.3	Amenazas	120
11.4	Debilidades	120
12	Objetivos estratégicos	120
13	Plan de marketing	121
13.1	Plan de marketing mix	121
14	Plan de Recursos Humanos	123
14.1	Organigrama	123
14.2	Análisis de los puestos de trabajo	123
14.3	Reclutamiento y selección de personal	124
14.4	Política de retribución	124
15	Plan Económico y Financiero	125
15.1	Plan de inversión	125
15.2	Plan de financiación	126
15.3	Explicación de la cuenta de Pérdidas y Ganancias	126
15.4	Desarrollo del balance de situación	129
15.5	Desarrollo del Cash Flow	130
15.6	Desarrollo del punto muerto	131
15.7	Análisis financiero	131
16	Gestión del proyecto	133
16.1	Planificación del proyecto	133
17	Conclusiones y trabajo futuro	134
17.1	Conclusiones	134
17.2	Evaluación del trabajo realizado	136
17.3	Trabajos Futuros	137
18	English abstract	138
18.1	Introduction	138

18.2	Company description	141
18.3	Strategic objectives	143
18.4	Marketing plan	144
18.5	Conclusions and future work	145
19.	Anexos	149
19.1	Bibliografía:	149
19.2	Glosario de términos y acrónimos.	154
19.3	Tablas anexas de cuentas.....	155

1. Introducción

El objetivo de este trabajo desarrollará una herramienta que sea capaz de automatizar el proceso de búsqueda de los términos y sus frecuencias de aparición en las páginas de Facebook que se dedican al negocio y/o la venta de productos alimenticios por internet.

La creciente importancia de Internet para el comercio electrónico, y en concreto de las redes sociales en la promoción de productos, hace necesario el desarrollo de herramientas que monitoricen la publicación de contenidos, para mejorar su difusión.

La naturaleza de buscadores y recursos es principalmente de carácter textual, por lo que la terminología empleada en ellos adquiere una especial relevancia. La terminología está, en este caso está centrada en el sector de la alimentación.

Se han seleccionado una serie de páginas de Facebook del sector de la alimentación. La razón para elegir Facebook es que es la red social con más usuarios, ya que cuenta con 1550 millones. Otras redes sociales como Youtube tiene 1000 millones, Twitter 350 millones e Instagram 400 millones [44].

A continuación se presenta un breve estado del arte sobre la optimización web, esto ayudará a entender el problema que se plantea resolver.

1.1 Planteamiento de la cuestión.

La explosión de uso de Internet desde principios de los años noventa ha causado que el contenido en línea se haya disparado desde entonces. Los crecimientos de contenido en Internet crecen exponencialmente.

Este crecimiento exponencial requiere de elementos de trabajo especializados ya que las herramientas genéricas cumplen satisfactoriamente los campos de trabajo generales. Pero ante el crecimiento de volumen de información en la web se necesitan herramientas más potentes.

No obstante no hay actualmente una respuesta concreta a cada uno de los distintos segmentos de la Web. En concreto para el creciente mercado de la venta online, el cual se encuentra en pleno auge. Este mercado de tiendas online también necesita de herramientas especializadas.

Debido al crecimiento del volumen de información se ha vuelto más difícil encontrar contenido útil. Por ello la Web introdujo la indexación, con el objetivo de etiquetar y ordenar el contenido como si del índice de un libro se tratase.

Esta es la idea principal de la que surgieron los motores de búsqueda. La tarea de estos es recorrer las distintas páginas para recopilar toda la información rastreada e incorporarla a los índices que mantienen.

Cabe mencionar que estos motores de búsqueda poseen algoritmos que les ayudan a ofrecer resultados a los usuarios que buscan contenidos en la web. Pero la clave reside también en cómo se ordenan estos contenidos según la relevancia que tengan para el cliente.

Esta relevancia se suele traducir en la popularidad de una página en concreto. Esta popularidad es la que asigna la ordenación de los distintos sitios web. Esta popularidad se calcula en base al

número de enlaces que recibe la página y la popularidad que tengan estos, lo cual corresponde al posicionamiento externo.

El posicionamiento web es clave para las páginas web de las empresas, ya que ante la cantidad de contenidos que hay en la Web, los usuarios leerán solo los primeros resultados quedando los demás ignorados por la mayoría de usuarios.

Es por ello que las técnicas y herramientas SEO ayudan a las páginas web nuevas y existentes a mejorar la estructura, la visibilidad y la calidad de su contenido para ascender en el ranking del posicionamiento web.

Actualmente en el mercado de las herramientas SEO se pueden encontrar muchísimas herramientas que ayudan a mejorar el posicionamiento de las páginas consultando diferentes parámetros y propiedades de la página consultada.

A pesar de que estas herramientas facilitan mucha información sobre las páginas consultadas como puede ser su “page Rank (no se publican en datos sobre su implementación en Google)”, “domain authority”, “Alexa Rank”, “términos clave”, número de enlaces entrantes (backlinks), presencia en las redes sociales y más información. Estas herramientas son genéricas y no se centran en ningún segmento en concreto.

No obstante algunas de estas herramientas no abarcan la principal red social del mercado que es Facebook. Ya que esta página requiere el uso de una API propia de pago para realizar visitas a los perfiles privados, debido a la política de privacidad de esta página. Permitiendo solamente el acceso a las páginas de interés común y al de las empresas.

Aunque también existe una forma de posicionamiento interno que depende en gran medida de la distribución y de cómo se presenta la información en la página. Para poder hacer de ésta una página más responsiva y con un contenido más claro que pueda ser leído fácilmente por los motores de búsqueda.

Sin embargo este sistema tiene un claro inconveniente ya que trata a la web como un elemento estático. Por lo que los sitios web que son importantes ahora lo serán aún más en el futuro. Por lo que los sitios nuevos lo tienen más difícil para alcanzar a los antiguos ya que en su clasificación gozan de menos popularidad.

Este nuevo sitio web puede contener información de calidad que ofrece a los usuarios de la web. No obstante al no ser más popular que los sitios antiguos se encuentra con la barrera de que solo unos pocos usuarios llegarán a conocerlo. Debido a esta visión estática de la web, las nuevas páginas están en clara desventaja con las antiguas.

1.2 Objetivos

Tras detallar el problema sobre las distintas herramientas SEO existentes y las desventajas de la visión estática de la web se procede a presentar una herramienta de exploración en Facebook. Esta herramienta permite al usuario explorar la frecuencia de aparición dentro de las distintas páginas del sector de la alimentación

Esta herramienta pretende ser una alternativa al resto de herramientas SEO actuales, ya que al ser de uso gratuito y específico para un ámbito concreto, que en este caso es la alimentación, puede servir al usuario para analizar un término clave a lo largo del tiempo. Aunque de momento solo se analiza la información relativa al año 2015.

Se ha comenzado con un dominio más específico como es el de la alimentación con el objetivo de crear la primera versión de la herramienta. No obstante se podría adaptar fácilmente a otros dominios si fuese requerido.

Se pretende por ello analizar las principales herramientas SEO del mercado para observar el comportamiento de éstas. De esta forma se podrán analizar las carencias y ventajas de las mismas frente a la aplicación propuesta.

Por consiguiente es necesario analizar las páginas que se dedican a la venta online de productos alimenticios a las cuales va dirigida esta herramienta ya que son un sector importante y muy interesado en este tipo de herramientas.

Lo que se pretende conseguir es una herramienta que sea capaz de ayudar a los usuarios a mejorar su posicionamiento SEO. Ya que por lo que se ha explicado antes para las páginas de reciente creación es difícil competir contra las páginas más antiguas.

1.3 Contenido de la memoria

Este punto detalla lo que se explica dentro de cada uno de los capítulos y anexos del documento.

En el primer capítulo bautizado como “Introducción” ofrece una visión general del proyecto, así como los objetivos marcados, y también el resto de contenidos en el documento.

El segundo capítulo es “El estado del arte”, el cual explica el contexto actual del tema del cual se va a tratar. En él se explica el algoritmo Pagerank que usa Google, otras herramientas para analizar la popularidad de las páginas, las tecnologías que se han usado para el desarrollo, explicación breve de técnicas de Google Hacking, una breve descripción del mercado de las ventas online de alimentos y tiendas con mejor valoración del año 2015.

El tercer capítulo es “Análisis de la solución” en el cual se hace una definición de como es el sistema, los requisitos de este y la identificación de los subsistemas que componen la aplicación.

El cuarto capítulo se llama “Diseño”, en el cual se detalla la arquitectura del sistema, la interfaz de usuario que se ha usado, identificar y especificar las clases que se han usado, diseño físico de las bases de datos y como se ha hecho la migración de datos inicial.

El quinto capítulo es “Especificación técnica del plan de pruebas, en el cual se incluyen las pruebas realizadas para comprobar el correcto funcionamiento del sistema.

El sexto capítulo se llama “Micro lenguaje del segmento de la alimentación”, en el cual se han analizado los términos relevantes en el ámbito de la venta de productos alimenticios en Internet.

El séptimo capítulo se llama “Taxonomía del micro lenguaje”, en el cual se lleva a cabo un estudio de las diferentes clasificaciones de los alimentos basándose en estándares europeos.

El octavo capítulo es la “Descripción de la empresa”, en el cual se detalla el perfil de la empresa y el proceso productivo que esta sigue.

El noveno capítulo es el “Análisis de la situación interna de la empresa”, donde se especifican las capacidades personales, técnicas, comerciales, financieras y de gestión de la empresa creada.

El décimo capítulo es el “Análisis de la situación externa de la empresa”, donde se ha hecho una análisis PEST para analizar algunas de las variables de entorno global. También se ha hecho el análisis de las 5 fuerzas de Porter para estudiar el entorno específico.

El undécimo capítulo es el “Análisis DAFO” en el cual se analizan las oportunidades, fortalezas, amenazas y debilidades de la empresa que se ha creado en este proyecto.

En duodécimo capítulo son los “Objetivos estratégicos”, en el que se definen las metas y estrategias para lograr los objetivos a largo plazo de la empresa.

El decimotercer capítulo es el “Plan de marketing”, en el cual se explican los planes de promoción del producto y la empresa en el presente y en el futuro.

El decimocuarto capítulo es el “Plan de recursos humanos”, donde se presenta el organigrama de la organización, el análisis de los puestos de trabajo, el reclutamiento de personal y la política de retribución.

El decimoquinto capítulo es el “Plan económico y financiero”, donde se detallan los planes de inversión, financiación, los flujos de caja, el punto muerto, la cuenta de pérdidas y ganancias y otros datos de interés para analizar económicamente la viabilidad de la empresa.

El decimosexto capítulo es la “Gestión del proyecto”, en el cual se detalla la planificación que se ha estimado y seguido para la realización de este proyecto.

El decimoséptimo capítulo son las “Conclusiones y Trabajo futuro”, donde se resume de manera breve el trabajo que se ha realizado y se dan recomendaciones para seguir desarrollándolo en el futuro.

El decimoctavo capítulo es el "English abstract" la cual resume brevemente algunos apartados de la memoria en lengua inglesa.

En la parte final hay varios anexos donde se completa la memoria dando datos que ayudan a la comprensión del documento. También se añaden términos y acrónimos especiales así como las fuentes citadas.

2. Estado del Arte

A continuación se explican las distintas aplicaciones y el material encontrado para desarrollar herramientas SEO. Las herramientas SEO (Search Engine Optimization) se basan en optimizar la estructura y la información de la página web, para mejorar la presentación de la página a los distintos buscadores.

Se van a analizar las distintas herramientas que existen actualmente en el mercado para poder sopesar sus ventajas y sus desventajas. También se analizará el algoritmo usado por Google para calcular el posicionamiento de cualquier página web.

Se explicarán en detalle algunas características de algunas páginas web de alimentación online que han sido destacadas por la asociación de comercio electrónico “club ecommerce awards España” [10]. No obstante se analizarán otras más que han sido destacadas en otros foros como se comentará más adelante.

Se detallarán minuciosamente las propiedades que tienen las distintas páginas, como son el juego de las redes sociales y los términos clave que estas usan. De estos términos clave se explicarán su relevancia en este contexto y la frecuencia con la que aparecen.

La aplicación se basa en el entorno Netbeans de programación Java, este entorno permite desarrollar con distintos sistemas gestores de bases de datos las aplicaciones que se pretenden desarrollar sobre este entorno.

2.2 El algoritmo PAGERANK.

El algoritmo Pagerank fue patentado por Google en el año 1999 para asignar de forma numérica la relevancia de los documentos o páginas web. Estos documentos eran ordenados gracias a un motor de búsqueda para determinar la relevancia de la página. [1]

Fue desarrollado por los fundadores de Google (Larry Page y Sergey Brin) usando una amplia estructura de enlaces para indicar el valor de un documento o página. Esto quiere decir que si una página en concreto apunta hacia otra, es decir, tiene un enlace hacia esa página. Se considera como un punto a favor de esa página.

Pero el razonamiento de Google iba más allá debido a que se analizaban también las fuentes, es decir, las páginas que contenían esos enlaces, si éstos provienen de sitios de confianza como son las páginas oficiales. Es por ello que se tienen en cuenta la suma del número de enlaces y la popularidad de los mismos. La idea es que este algoritmo representase la importancia de estas páginas. [1]

$$PR(A) = (1 - d) + d \sum_{i=1}^n \frac{PR(i)}{C(i)}$$

Ilustración 1. PAGERANK Fuente [1]

Donde las variables son:

- $PR(A)$, es el PAGERANK de la página A.
- $PR(i)$, son los valores de PAGERANK que tienen cada una de las páginas i que enlazan a A.
- $C(i)$, es el número de enlaces salientes de la página i . (sean o no hacia A)
- d es un factor de amortiguación que tiene un valor entre 0 y 1

(Fuente Wikipedia [1])

No obstante al ser el número de enlaces una medida objetiva para este algoritmo comenzó la manipulación del algoritmo para conseguir más enlaces de páginas con un Pagerank alto. De esta forma se conseguía posicionar mejor las páginas que interesaban a los defraudadores.

El Pagerank era actualizado para subir o bajar la calificación de un sitio web, posteriormente se hacía cada tres meses, después seis y por último anualmente. La manipulación por parte de

Webmasters y expertos SEO, llevó a Google a abandonar paulatinamente algunas herramientas que se ofrecían a partir de este algoritmo como era Toolbar Pagerank [46]. No obstante desde el año 2013 Google ha decidido no dar más información acerca de este algoritmo para no dar facilidades a los manipuladores para ejercer su engaño [2].

La degradación del Pagerank ha abierto nuevas puertas a nuevas variables para influir en el posicionamiento como pueden ser la autoridad de dominio, al análisis de palabras clave [47] y el algoritmo MOZRank.

2.2 Otras herramientas para analizar la popularidad

Se van a detallar algunas técnicas y algoritmos alternativos al clásico Pagerank de Google. Estas sirven para calcular la popularidad de las páginas web.

2.2.1 Análisis del lenguaje

Para poder posicionar correctamente una página en Google u otros motores de búsqueda como Yahoo o Yandex, se deben escoger minuciosamente las palabras que se utilizan en los títulos, el texto y los encabezados. Usar palabras extrañas o faltas de ortografía pueden empeorar el posicionamiento ya que los robots que usan los motores de búsqueda para indexar no localizarán correctamente una página.

Por otro lado usar palabras demasiado comunes como por ejemplo “Agua”, “Noticias”, “Jamón”, etc. Pueden dificultar la tarea de posicionamiento ya que muchas páginas usan estos términos y por lo tanto la página a posicionar entrará en esta feroz competencia por hacerse con los primeros puestos del ranking.

Debido a este esfuerzo de tiempo y dinero se debe emplear de forma correcta los términos que se utilizan en la página para dar al lector una idea rápida sobre el tema que se va a tratar en el texto. Repetir con frecuencia los términos clave para poder dejar claro la idea del texto.

Estos términos clave o keywords ayudan a los usuarios a encontrar la página deseada, ya que estas serán las que el usuario vaya a emplear en el navegador. Esta tarea de búsqueda debe culminar encontrando los términos que sean más rentables para el usuario por lo que la página debe tratar de captar la mayor cantidad de términos clave que pueda. [4]

Es preciso detectar también las palabras clave que apenas tienen búsquedas, o son poco usadas por los usuarios, ya que aportarán pocas visitas a la página. Algunas de estas palabras clave varían con el uso del tiempo, ya que en función de la época del año, se usan más unas que otras. Por lo que es preciso conocer la estacionalidad de éstas para poder dirigir campañas de publicidad o posicionar una página.

En el uso de palabras clave es común que se usen también combinaciones de dos o tres palabras para fabricar una sentencia clave que equivaldría a lo que es una palabra clave. No obstante es posible encontrar combinaciones de cadenas más largas de texto.

Para encontrar estas palabras clave es vital centrarse en un tema en concreto para localizar correctamente el tema del que se habla o los productos que se van a ofrecer [4]. Por ello estas frases deben distinguir a la página de forma que los usuarios que hagan la búsqueda estén buscando algo relacionado con el tema de la página.

2.2.2 MOZRANK

Mozrank surge como una alternativa de PageRank, a diferencia de éste se aplica a todas las páginas de internet. Emplea números fraccionarios para clasificar las diferentes páginas. Usa la escala logarítmica del 1 al 10. Siendo 1 la calificación más baja y 10 la más alta.

La calificación de la página se mejora en función de la calidad de los enlaces que apuntan a la página, ya que a mayor calidad, mejor calificación obtendrá la página. Mozrank solo tiene en cuenta un link por cada dominio. [3]

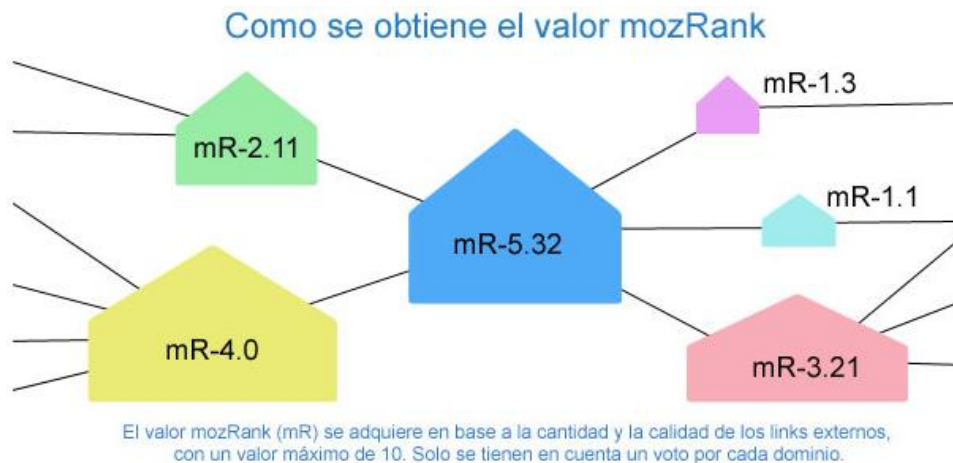


Ilustración 2. MOZRANK. Fuente [3]

Para ello el sistema se apoya en Linkscape que contiene una gran base de datos sobre los links que hay entre las distintas páginas web. Éste se actualiza periódicamente con los links que hay entre las páginas.

2.2.3 Ahrefs

Se trata de otro servicio que usa un ranking en función del número de backlinks de una URL y la popularidad de éstas. Utiliza una métrica entre 1 y 100. Por lo que tiene en cuenta la popularidad del enlace, distinguiendo si es de calidad o no [5].

2.2.4 Índice Alexa

Analiza la popularidad del sitio web y el tráfico que ésta recibe. La popularidad del sitio es una combinación de visitantes únicos y número de páginas visitadas (URL distintas). Las visitas se tienen en cuenta cada día para analizar el tráfico recibido [5].

2.3 Tecnologías necesarias para el desarrollo

En el siguiente apartado se comentan las distintas tecnologías utilizadas para llevar a cabo el desarrollo del proyecto. Se comentará brevemente las razones de la elección de estas tecnologías frente a otras.

La elección de las distintas tecnologías se hace en función de los criterios de:

- **Coste:** dependiendo del coste de adquirir las licencias de uso de la tecnología.
- **Rendimiento:** medida con la que en función de un determinado tiempo, se obtiene un mejor resultado.
- **Experiencia:** haber trabajado anteriormente con las tecnologías citadas.
- **Licencia:** si tiene licencia de uso libre o requiere algún pago por derechos de uso.

2.3.1 Lenguaje de programación

Los lenguajes de programación que se han usado para programar este sistema han sido:

- Java
- HTML
- XML
- SQL

Estos lenguajes se han usado debido a la alta experiencia del desarrollador con todos ellos, y a estar más familiarizado con los entornos en los que se usa. El coste de uso de estos lenguajes es 0 ya que son gratuitos para cualquier usuario.

Cuentan con el soporte de muchísimos entornos de programación y multitud d APIs para poder operar. Lo cual les hace muy atractivos para usar. El rendimiento que se puede sacar con estos lenguajes es alto ya que son muy flexibles y poseen un amplio espectro de herramientas y ayuda en internet para poder programar con ellos adecuadamente.

2.3.2 Sistema operativo

El sistema operativo sobre el cual se ha llevado el desarrollo es Windows 10, éste sistema operativo tiene un alto rendimiento [48]. El desarrollador posee una alta experiencia con este sistema operativo por lo que juega un papel muy importante en la decisión de seleccionar a este frente a otros.

El tema del coste no es muy relevante ya que la UC3M facilita este software a sus alumnos. Por lo que ante esta ventaja se decide seleccionar este sistema operativo frente a los otros que pueden ser Linux o Mac OS.

El inconveniente de Linux y de Mac es la poca experiencia del usuario con estos sistemas por lo que se ha decidido escoger Windows 10.

2.3.3 Entorno de programación

La plataforma escogida para ayudar a programar sobre los distintos lenguajes que se han usado en el desarrollo es Netbeans. Este entorno de programación ofrece múltiples ayudas y aplicaciones con lo cual se integra perfectamente para un desarrollo web de estas características.

El desarrollador posee amplia experiencia con este entorno de desarrollo lo que le da mucho rendimiento para desarrollar. La gran ventaja de este entorno es que viene integrado dentro de otra aplicación más grande que se llama XAMPP. Esta aplicación contiene más funcionalidades pensadas para el desarrollo web lo que facilita enormemente la tarea.

Otros entornos como son Eclipse o Microsoft Visual Studio 2010, también ofrecen funcionalidades parecidas. En el caso de Eclipse el desarrollador posee asimismo gran experiencia y es gratuito. No obstante no ofrece las mismas facilidades que Netbeans.

Microsoft Visual Studio 2010, tiene licencia de uso, no obstante es gratuita gracias a un convenio de la UC3M, con Microsft. No obstante el desarrollador no posee tanta experiencia con este entorno de desarrollo.

2.3.4 Sistema gestor de bases de datos

El sistema gestor de bases de datos se encarga de controlar y administrar la base de datos a la que accede la aplicación. El sistema que se ha usado para el desarrollo es MariaDB, el cual es un desarrollo de MYSQL, que viene integrado con el sistema XAMPP.

El sistema MariaDB es sencillo de usar y viene con herramientas de ayuda para su gestión como phpmyadmin. El usuario posee más experiencia y conoce bien esta tecnología. El coste es gratuito lo que la hace aún más preferible.

Otras opciones de uso serían Access, Microsoft SQL Server y Oracle 2003. No obstante el desarrollador opta por MariaDB que es gratuita y sencilla de usar.

2.3.5 Gestión del proyecto

Se analizan las distintas herramientas utilizadas para la debida gestión del proyecto. Estas herramientas incluyen las tareas de documentación, gestión de incidencias u otros eventos y la planificación de tiempos.

Planificación de Tiempos

Para llevar un registro de los tiempos y poder programar las tareas ordenadamente se tenía en mente utilizar los Gantt-Chart que están disponible con la herramienta draw.io que es gratuita y no requiere instalación.

Documentación

La gestión de la documentación se ha llevado a cabo en el Office Word de Microsoft ya que esta herramienta es muy potente y admite mucho contenido de distinto tipo, no solo texto. Existen otras herramientas de carácter más centralizado como Enterprise Architecture o Microsoft Visual. No obstante las necesidades del proyecto no requieren de herramientas tan sofisticadas. También se ha recurrido a otras herramientas de Office como son Excel para los gráficos y las tablas y Power Point para diapositivas y figuras que se utilizan para crear esta memoria

Registro de incidencias

Para la gestión de incidencias se podría haber usado Microsft Visual Studio no obstante por la falta de tiempo y experiencia en este campo se van a documentar las incidencias en el mismo documento Word en el que se escribe la memoria para ahorrar recursos.

Diagramas UML

Se ha seleccionado la herramienta draw.io para llevar a cabo la gestión de los diagramas UML. Esta herramienta es gratuita y no requiere instalación, su uso es muy sencillo por lo que aumenta enormemente el rendimiento debido a la necesidad baja de conocimientos previos para usarla.

2.3.6 Gestión de pruebas

La comprobación de la fiabilidad de un sistema es una parte fundamental en el desarrollo de la aplicación como comprobación de que el sistema hace lo que tiene que hacer. Esto significa que la aplicación necesita de una pila de pruebas completa para comprobar las distintas funcionalidades del sistema.

Cabe destacar que la pila de pruebas no comprueba perfectamente todos los campos de la aplicación ya que asegurar que todo funcione perfectamente es imposible. No obstante se debe certificar unos servicios básicos para poner el sistema en producción

Las pruebas unitarias, funcionales y no funcionales se van a gestionar en Word ya que usar otras herramientas más sofisticadas como NUnit de Microsoft Visual Studio escapan al alcance del desarrollador por la poca experiencia que tiene con ella.

2.3.7 Metodología de desarrollo

La metodología de desarrollo que se ha llevado a cabo para la gestión de este proyecto es SCRUM. Se ha seleccionado esta metodología debido a que es la que se ha enseñado en la carrera con mayor énfasis durante los últimos años.

Es una metodología ágil ideal para los proyectos complejos donde se necesita obtener resultados pronto, donde los requisitos son cambiantes o poco definidos, donde la innovación, la competitividad, la flexibilidad y la productividad son fundamentales [8].

El proceso de SCRUM se lleva a cabo en periodos de tiempo cortos y fijos (Iteraciones), habitualmente cada 2 semanas. Cada una de las iteraciones debe proporcionar un resultado entero, ésto hace que incremente el desarrollo del producto final [8].

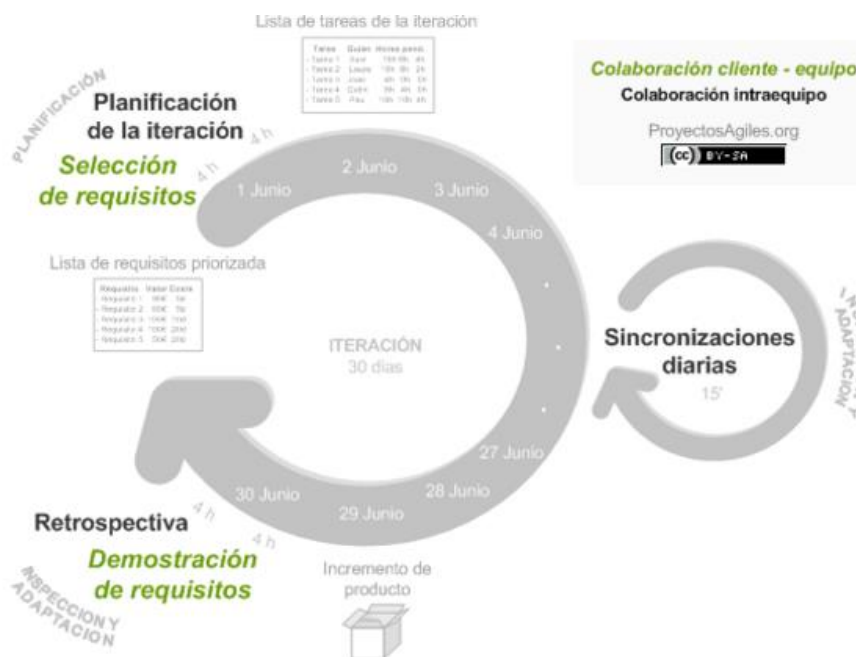


Ilustración 3. Fuente [8]

Ilustración 3. Scrum iteración. Fuente [8]

Las características cambiantes de este proyecto y su complejidad han conducido a que la mejor elección de metodología sea SCRUM. Ya que el desarrollo en paralelo de las distintas actividades y desarrollo de la aplicación requieren de prácticas que con el mínimo esfuerzo maximicen los resultados.

La gestión del cambio es algo que está a la orden del día en este proyecto por lo que se necesita de una metodología que pueda acatar estos cambios. Los procesos iterativos ayudan a que estos cambios no afecten a otros módulos, y en el caso de que se vean afectados que sean los menores posibles.

	CMM	ASD	Crystal	DSDM	FDD	LD	Scrum	XP
Sistema como algo cambiante	1	5	4	3	3	4	5	5
Colaboración	2	5	5	4	4	4	5	5
Características Metodología (CM)								
-Resultados	2	5	5	4	4	4	5	5
-Simplicidad	1	4	4	3	5	3	5	5
-Adaptabilidad	2	5	5	3	3	4	4	3
-Excelencia técnica	4	3	3	4	4	4	3	4
-Prácticas de colaboración	2	5	5	4	3	3	4	5
Media CM	2.2	4.4	4.4	3.6	3.8	3.6	4.2	4.4
Media Total	1.7	4.8	4.5	3.6	3.6	3.9	4.7	4.8

Tabla 1. Ranking de "agilidad" (Los valores más altos representan una mayor agilidad)

Ilustración 4. Ejemplo de metodologías ágiles. Fuente [9]

Cabe mencionar que no se va a seguir la metodología SCRUM estrictamente, sino que se van a hacer estimaciones de esfuerzo y desarrollo. Éstas son siempre alterables a lo largo del proyecto por lo que al principio se hará más esfuerzo en recopilar información y datos para después hacer más énfasis en detallar los requisitos que, en el caso de este proyecto, han sido ligeramente modificados al principio del desarrollo.

2.3.8 Weka

Weka es una herramienta software para el aprendizaje automático y la minería de datos. Contiene una serie de algoritmos para analizar los datos que se le introducen a través de unos archivos de texto con un formato especial para ello. Este formato es el .ARFF, con el que se ejecutan las pruebas para comprobar la efectividad del análisis.

Esta herramienta ofrece un entorno sencillo de usar con información acerca de distintos parámetros, que se pueden analizar. Estos parámetros y las variables de salida son personalizables lo que le da a la herramienta una gran versatilidad. El desarrollador ha usado esta herramienta durante la carrera por lo que tiene experiencia en el uso de la misma.

2.3.9 Resumen de las tecnologías

Se resumen a continuación las tecnologías que se han usado para desarrollar este proyecto de forma breve y sencilla.

Gestión del proyecto	Draw.io Microsoft Office
Gestión de pruebas	Microsoft Office
Metodología de desarrollo	Metodología ágil
Sistema Operativo	Windows 10
Entorno de programación	Netbeans
Sistema gestor de bases de datos	MariaDB
Lenguaje de programación	Java, SQL, HTML, XML

Tabla 1 Resumen de tecnologías

2.4 Google Hacking:

Para llevar a cabo una mejor exploración se han ejecutado una serie de consultas de Google Hacking, Las cuáles se muestran a continuación para dar una idea general del método a seguir para encontrar páginas relacionadas en el sector de la alimentación y encontrar así páginas relevantes y analizar la calidad de las mismas.

Para encontrar estas instrucciones se han seguido diferentes tutoriales o manuales. (Fuentes [14], [15]). Estas instrucciones pueden aportar información sobre cómo indexa Google la información acerca de términos y cómo de relevantes son.

- Otra más interesante para llevar a cabo un crawling pasivo sería ejecutar esta sentencia:

Site:www.microsoft.com Bill Gates

Lo que va a hacer es ejecutar una búsqueda dentro de las páginas de Microsoft que contengan la cadena de texto de Bill Gates.

Site:www.cerezasdelbierzo.com Vino

Lo mismo se puede hacer con esta sentencia para encontrar páginas relevantes de la página (www.cerezasdelbierzo.com) con información acerca del vino.

- Otra instrucción muy importante es “allinurl” la cuál nos va a dar todas las páginas que contengan la cadena de caracteres que pongamos a continuación. Por ejemplo, si queremos visitar la lista de catálogos de una tienda online debemos escribir:

Allinurl:saboresibericos.com/collections

De esta forma se sacan todas las páginas que contengan la palabra “collections” en la Url. Se pueden ejecutar una gran variedad de instrucciones como por ejemplo sacar la lista de páginas que utilizan las palabras clave como oferta, descuento y fresco.

- Otra sentencia muy parecida es “allintitle” con esta instrucción se pueden sacar todas las páginas con unas determinadas palabras de interés como pueden ser supermercado o tienda online.

Allintitle:tienda online

- Para poder ver información de un determinado portal se usa el operador “info” seguido de la dirección de la página de la que se quiere sacar la información. Un ejemplo sería: (info:Carrefour.es)
- Para poder consultar la información que google tiene en caché sobre una determinada página se puede hacer con el operador “cache”

Cache:saboresibericos.com

Aparecerá la información referente a la página que se haya seleccionado con un mensaje en la parte de arriba como se muestra a continuación.



Ilustración 5. Operador google en SI. Fuente propia

- El operador “filetype” busca el tipo de archivo que se le indica, se ha de escribir la extensión del archivo que se quiere buscar. Ofrece una lista de todos los archivos del tipo especificado con la palabra clave tanto en la URL como en el título del texto que se añade a continuación de la extensión del archivo. Un ejemplo sería:

Filetype:pdf recetas

- El operador “define” devolverá la definición de la palabra clave que se introduce. Se pueden definir muchos términos como por ejemplo tienda, fruta, venta o producto. No obstante va a devolver muchos enlaces que pueden definir distintos ámbitos de uso de la palabra

Define:producto

- Otro operador muy importante para buscar entre dos rangos de números es los dos puntos seguidos “..” que nos permitirá sacar resultados que contengan alguna consulta que está comprendida entre los dos términos que vamos a especificar. Un ejemplo de esto sería:

Ventas jamon 2012 .. 2015

- Para poder ver el dominio quitando alguna página que no nos interese ver podemos escribir:
Site:saboresibericos.com –site:www.saboresibericos.com/collections

- Una operación muy curiosa para ver los puertos abiertos de una determinada página puede ser el siguiente operador. Éste operador encontrará los puertos abiertos del servidor que tenga un dominio “eu” con el puerto 10000 abierto.

site:eu inurl:10000 webmin

- Se pueden encontrar muchos más operadores en páginas como ésta:
<http://thenewartofwar.eu/google-hacking/>
- Un operador muy importante devuelve las imágenes que se hayan sacado de una cámara digital o desde un teléfono móvil. Esto es debido a que muchos de estos dispositivos guardan las imágenes que toman en una carpeta con este nombre Lo cual puede ser muy interesante para recopilar información acerca de las fotos que las otras páginas competidoras tienen en su página o han guardado para hacer publicidad.

Index.of.dcm

2.5 El mercado de las ventas online de alimentos

El mercado de la venta de productos alimenticios está en auge en España, esto es debido a la alta penetración de los Smartphones en el mercado español. La predisposición de los españoles para comprar por internet es alta según indica la asociación de Ecommerce en España. [16]

El 0,6% en el total de las ventas del mercado de gran consumo fue en el segmento de la alimentación en el año 2013. Según la consultora Nielsen la distribución online de alimentos en octubre de 2014 fue de 24 millones de euros. [16]

En el año 2015 la cifra aumentó significativamente ya que cada vez más hogares españoles van incorporando la costumbre de hacer la compra por Internet. Esto lo demuestra el hecho de que las ventas desde junio 2014 hasta 2015, donde crecieron un 25% según los datos de la consultora Nielsen. [17]

Esta consultora también ha observado que hay un notable aumento de las ventas online en todas las categorías y en especial en las que tradicionalmente eran más adversas a este mercado. Estas categorías eran los refrigerados o la charcutería que crecieron un 15% aproximadamente. [17]

El consumidor habitual suele ser un hogar de uno o dos miembros con hijos. En general suelen ser de clase media-alta, con preferencia a comprar a principio de mes, cuando se ha recibido la nómina. El crecimiento de estos productos se puede ver en el siguiente gráfico:

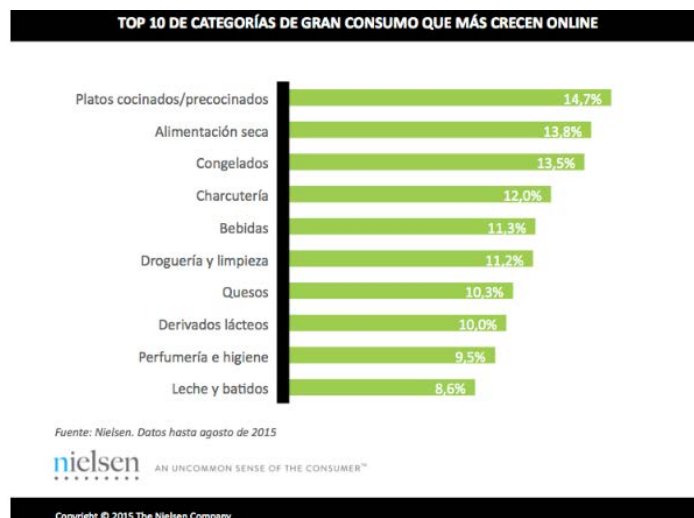


Ilustración 6. Categorías de consumo online. Fuente [17]

Como se puede observar la alimentación que es perecedera a largo plazo sigue siendo la que más se consume. El triunfador por excelencia son los platos precocinados. Las compras se suelen hacer en días laborables por lo que se asume que la flexibilidad de este sistema de venta online supera a la del método de compra física tradicional.

En el método de compra tradicional el consumidor suele comprar durante el fin de semana, mientras que como se ha explicado anteriormente la venta online se suele hacer en días laborables.

No obstante esta venta online no acaba de despegar del todo ya que en países como el Reino Unido y Estados Unidos estas cifras son mucho mayores (1026 millones de libras en el caso del Reino Unido).

Ello es en parte debido a que las grandes cadenas de supermercados no han mostrado un gran interés por este sistema de ventas. Además de que una parte de las ventas responde a una compra impulsiva, cosa que se con menor frecuencia por Internet. [18]

Las empresas del sector en gran medida no han dado con una buena fórmula para hacer de este negocio algo muy rentable, ya que le venta online resulta cara (entre 6 y 12 €). Este coste es demasiado alto para un sector con márgenes de beneficio bajos. Esto es solo el 1% del total del negocio por lo que no suscita el interés de las grandes superficies. [18]

El problema en mayor medida es la logística ya que se necesita una fuerte inversión en los canales de distribución a la vez que se necesita que se distribuya de una forma barata y eficiente. Es por ello que empresas como tudespena.com o Ulalbox.com opten por externalizar sus servicios de distribución. [18]

El gran objetivo de estas empresas es conseguir distribuir los productos frescos como se ha conseguido en el año 2015, ya que esta barrera era la que impedía crecer y fidelizar al cliente. La facturación del mercado de la venta online de alimentos es de 260 millones. Esta cifra es baja si se considera contra el total del segmento.

Si se compara el volumen de ventas de este segmento contra los demás segmentos de venta online como es la compra de viajes y entradas de cine, lo cierto es que el porcentaje es realmente bajo, ocupando las últimas posiciones. Se puede observar en el grafico a continuación:

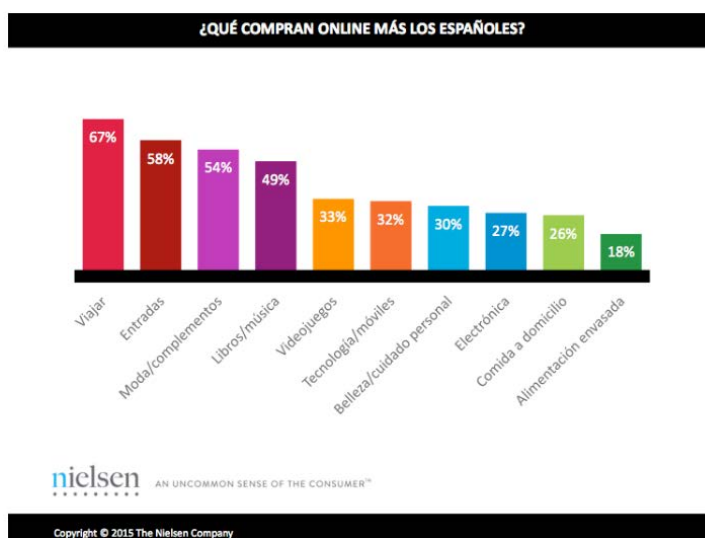


Ilustración 7. Compradores online franjas. Fuente [19]

No obstante, cómo se ha mencionado en los anteriores párrafos, este mercado está en plena expansión gracias a los dispositivos móviles y la mejora de las tecnologías de comunicaciones. La digitalización de cada vez más procesos de compra y la adaptación de la sociedad actual a estas nuevas tecnologías de compra hacen cada vez más atractivo este modo de compra.

El usuario medio de estos servicios suelen ser profesionales en torno a los 30 a 45 años de edad que viven en hogares de 1 o dos personas. Perciben rentas altas y no tienen hijos por lo que prefieren emplear su tiempo en otras actividades de ocio.

2.6 Investigación sobre las mejores tiendas del año 2015.

Se ha procedido a investigar las mejores tiendas online en el año 2015 en la rama de alimentación y delicatessen. Se ha escogido como referente a la organización club Ecommerce [10] con reconocimiento internacional en el campo de las páginas y las empresas de venta de productos online. Esta página cataloga las diferentes páginas según la calidad de sus servicios y la calidad del portal web que emplean para mostrar al cliente.

Es importante recalcar que se han recogido los siguientes términos para hacer la búsqueda y filtrar con estas palabras las posibles preferencias de los clientes. Estas palabras se han escogido por su peso en las búsquedas de internet, todas ellas son palabras muy usadas por el público que suele buscar productos alimenticios en Internet. La recomendación de estas palabras vino de la herramienta Traffic Travis.

Artesanal, chorizo, jamón, comida, tienda, tienda de alimentación, comprar comida online, venta online, comida exclusiva, supermercado.

A continuación vamos a estudiar algunos de los portales relevantes en la comercialización de estos productos, según su orden de aparición en las consultas en el buscador.

2.6.1 Cerezas del Bierzo (<http://www.cerezasdelbierzo.com/>)

En primer lugar aparece “Cerezas del Bierzo”, esta página se dedica a la venta de todo tipo de productos alimenticios típicos de la región del Bierzo. Esta variedad abarca fruta fresca, embutidos, vinos, carne, conservas, cervezas artesanas, infusiones y platos (también bandejas) de pizarra.

Se han analizado algunas de las páginas de esta tienda online y se ha comprobado que cumple la mayoría de las recomendaciones SEO. No obstante la densidad de términos no es muy alta ya que apenas llega al 2,3 % de densidad en la repetición de términos clave.

El “PAGE AUTHORITY” de esta página es relativamente alto ya que es un 34,58, lo cuál la sitúa en una posición aventajada. El rango Alexa (Alexa Rank) la sitúa en el 5.619.038. Por lo que la página está bien situada en Google.

2.6.1.1 Términos clave

Estos términos se han explorado con la herramienta de traffic travis. Se observa que el término más repetido en la página es “bierzo”, seguido de vinos, platos, regalos gourmet, tienda gourmet, venta online de vino, venta aceite, venta de queso, venta de jamón, lotes y cerezas.

La densidad de términos clave no es demasiado alta, se encuentra en el 2,3% que se sitúa en la media de lo que se ha observado hasta ahora. Aunque este porcentaje es sobre unos 5 términos. Ello es debido a que las páginas de venta online se basan en un entorno gráfico más que en el texto descriptivo.

Otros términos formados por más de una palabra son bierzo lotes, en conserva, platos de pizarra y otros tantos más. No obstante estas palabras no están en los headers de las páginas y pocos de ellos suelen encontrarse en los títulos de las páginas. No obstante donde si suelen encontrarse es en la descripción de la página.

2.6.1.2 Redes sociales

La presencia en las redes sociales no es especialmente alta, ya que básicamente se concentra en google + y Facebook donde mantiene una actividad más fuerte. Esta actividad en Facebook se resume en unos 50 comentarios, otros 590 me gustas y compartido 870 veces. Por lo que se ha concentrado prácticamente toda la actividad en Facebook.

2.6.1.3 Enlaces

Esta página cuenta con un gran número de links internos en cada página, es decir, que hay un gran número de links que apuntan a páginas que están dentro del mismo dominio.

El número de backlinks de la página es alto ya que cuenta con 2494 a día 21 de marzo de 2016. Estos backlinks proceden de sitios como el periódico “El mundo” (tiene un alto “PAGE AUTHORITY”, cuanto más alto sea el “PAGE AUTHORITY” más fácil será que la página suba de rango).

Otras páginas como comidasmaitemontes.com que es otra página que ofrece servicios de catering con un alto tráfico. También aparecen en otras páginas así como varios foros sobre restaurantes y recetas como son leonenred.com y nextroom1.blogspot.com.es.

Incluso está referenciado en otras páginas como agroboca.com que es una tienda online de frutas y hortalizas. Esta página es apuntada por muchas otras que le han ayudado a escalar posiciones. Por lo que se ha comprobado, la mayoría de estas páginas son activas. Lo que le da aún más importancia a los enlaces que recibe de ellas.

2.6.2 LaTiendaHero (<https://www.latiendahero.es/>)

Esta tienda online está especializada en productos para bebés aunque ofrece también productos para toda la familia. En este caso pesa más en el estudio la alimentación para toda la familia por lo que la parte de bebés no se analizará.

En esta tienda se pueden encontrar productos como mermeladas, barritas de cereales, zumos, salsas y comida enlatada. Esta página posee una densidad de términos ligeramente más alta que la de otras páginas (2%), no obstante esta ligera diferencia no es un gran aporte para su definitivo posicionamiento.

El "PAGE AUTHORITY" de la página es de 22.44, lo cual es alto pero no está entre los valores altos que se han encontrado en otras páginas del mercado. No obstante el rango Alexa (Alexa Traffic Rank) es de 1.437.703. Por lo que se puede deducir que el tráfico de la página es elevado dentro de su categoría.

2.6.2.1 Términos clave

En esta página no hay una alta concentración de términos clave, ya que en los resultados arrojados en el análisis, la densidad ronda el 2%. No obstante en términos más importantes cabe destacar el propio nombre de Hero y tienda Hero que si está en el título y la descripción de la página con una densidad de aparición relevante.

Otros términos relevantes que se han encontrado son: tienda online, alimentación infantil, leches infantiles, postres, frutas y cereales infantiles. Lo que denota la especialización de esta tienda en productos infantiles.

2.6.2.2 Redes sociales

Esta página no tiene prácticamente presencia en las redes sociales más importantes como son Facebook, Twitter e Instagram. Ya que no se han encontrado tampoco enlaces relevantes desde éstas por lo que su presencia se puede decir que es muy baja.

2.6.2.3 Enlaces

La mayoría de enlaces que recibe esta página son del propio blog que tiene la compañía donde se publican noticias de interés como su email de contacto, cursos y consejos de alimentación. Este blog es su principal fuente de enlaces, que además cuenta con un page Rank notable (30%) lo que le aporta valor para ser posicionado. También aparece en algunos portales que hablan sobre el comercio electrónico como es por ejemplo cecarm.com, no obstante no son muy conocidos y apenas hacen referencia a esta página.

2.6.3 Chuchesonline (<https://www.chuchesonline.com/es/>)

Esta página se dedica al negocio de la venta de golosinas, tartas, chocolates, galletas, frutos secos, cromos, juguetes y artículos para fiestas y eventos. La página está especializada en chucherías, no obstante abarca un gran abanico de oferta de productos, que llegan desde máquinas para fabricar algodón de azúcar y piñatas para los niños.

Esta página cumple la mayoría de las recomendaciones SEO como son que el texto no tenga más de 100 palabras, que los links que tiene la página sean activos, la URL es correcta (la longitud es adecuada) y los títulos tienen la longitud adecuada. También las metadescripciones están puestas. Aunque las imágenes no tienen descripción cumple con la mayoría de los requerimientos para un buen SEO.

El "PAGE AUTHORITY" de la página es bajo (1.00), lo cual no ayuda para su posicionamiento en Google. Su "Alexa Rank" es 592.808. Este resultado no es especialmente alto frente al resto de páginas analizadas.

2.6.3.1 Términos clave

La densidad de los términos clave no es muy alta ya que, aparte de las preposiciones, las palabras que más destacan son: cookies, chuchesonline, chuches, chocolates, comprar chuches online y tienda. Pero éstas no se repiten mucho en el texto por lo que no se puede afirmar que su posicionamiento se base en la correcta colocación de los términos.

Éstos términos se concentran en torno al nombre de la marca y la especialización de ésta en el ámbito de las chucherías. No obstante hace poca referencia a los términos tienda y alimentación.

2.6.3.2 Redes sociales

Esta página tiene muy poca presencia en las redes sociales ya que no tiene ningún enlace de las mismas. No se ha promocionado en ninguna de las principales redes sociales (Facebook, Google+, Twitter, Pinterest). Solo tiene un enlace en LinkedIn.

Por lo que se puede concluir que no hay una presencia relevante en las principales redes sociales que promocionen la marca.

2.6.3.3 Enlaces

Los enlaces en valores absolutos de esta página no son altos ya que apenas tiene 11 enlaces internos a día 23 -4-16 y solo 7 enlaces externos. Estos enlaces vienen de una página de calificación de comercio electrónico (<http://marketing4ecommerce.net/chuchesonline-analisis-valoracion-y-comentarios/>). Esta página se dedica a analizar las distintas páginas y tecnologías que hay en el mercado.

No obstante este foro no tiene un "PAGE AUTHORITY" alto ya que solo tiene 9.28 sobre 100 puntos lo que no ayuda mucho al posicionamiento de la página. Sumado a esto la poca cantidad de estos enlaces que se han detectado no dan un alto desempeño para posicionar la página.

2.6.4 FrutasZabalza (<http://www.frutaszabalza.com/>)

Esta página se dedica a un amplio espectro de venta de productos alimenticios. El nombre de la página le viene dado de un municipio en Navarra llamada Zabalza. La variedad de productos que ofrece la página son:

Verduras, frutas, envasados (para congelador), pastas, huevos, lácteos (yogures y quesos), pescado, carne, legumbres, conservas, fritos, productos precocinados, miel y frutos secos, hongos (setas y níscalos), conservas y productos deshidratados.

La página contiene solo algunas de las recomendaciones SEO como incluir las descripciones de las imágenes o incluir las metadescripciones y tener unos títulos descriptivos. No obstante, la densidad de los términos repetidos no es muy alta, ya que apenas pasan del 2%. La Url y el título de la página no son muy descriptivos ya que hay otros negocios con nombres semejantes y esto puede llevar a confusión fácilmente.

El rango Alexa (Alexa Traffic Rank) es de 8.269.350, este rango es muy alto, ya que de las páginas estudiadas es de los más grandes que se ha encontrado. El "PAGE AUTHORITY" de la página es 32.77, un rango muy alto en comparación con las otras páginas estudiadas. Como se puede observar el tráfico de esta página es muy elevado lo que ayudará a posicionarla en Google.

2.6.4.1 Términos clave

La densidad de los términos clave de esta página es baja ya que apenas llegan al 3% en el mejor de los casos. Los términos clave más importantes que se han encontrado son: verduras, frutas, magdalena, comprar, Navarra, envasados (junto con frutas envasados, envasados pastas mary, envasados frutas verduras), cestas (verduras, envasados, frutas y huevos) y huevos.

Estos términos sí que ayudan a posicionar la página ya que es donde la empresa quiere destacarse, que es en la venta de productos frescos y de calidad. Aunque vende también productos envasados y deshidratados, los cuales también aparecen destacados en los términos clave de la página que más se repiten.

2.6.4.2 Redes sociales

La página tiene una actividad muy dispersa en las redes sociales, ya que aunque no destaca especialmente en ninguna de las redes sociales favoritas, abarca la mayoría de las redes sociales más importantes como son Facebook, LinkedIn, Google+ y Pinterest.

En Facebook tiene más repercusiones ya que ha sido compartida más veces su página web y tiene un me gusta. En Google+ y en Pinterest también tiene una presencia pequeña de 3 y 4 apariciones.

2.6.4.3 Enlaces

Esta página tiene un número significativo de enlaces ya que tiene a día 24/3/16 un total de 69 enlaces internos debido a su alta gama de productos. Tiene 7 enlaces externos hacia otras páginas. No obstante cuenta con 183 backlinks que le ayudan a escalar en Google ya que algunos de estos enlaces son de páginas oficiales con un "PAGE AUTHORITY" de más del 30%.

Entre estos enlaces se cuentan algunos relevantes como el periódico "El País", algunos blogs de recetas reconocidos, una web de marketing de Navarra especializada en el ámbito local, foros de turismo (nunet.com) e interés culinario (lasrecetasdemarichuyasmias.blogspot.fr) y otros foros de actividades deportivas.

Al ser la mayoría de páginas muy activas hace que la página pueda escalar en el posicionamiento global, lo que le da una gran ventaja frente a sus competidores. Al ser considerados estos enlaces de más calidad frente a otros que tengan un "PAGE AUTHORITY" más bajo (los que rondan el 3%).

2.6.5 Tienda Osborne (<https://www.tiendaosborne.es/>)

Esta tienda se dedica a vender jamón, lomo, vino, quesos, licores y packs que combinan su catálogo de productos. La tienda posee productos de diferentes categorías y precios dentro de cada rama de productos.

La página cumple con la mayoría de las recomendaciones SEO ya que cuenta con títulos adecuados que describen correctamente la página. También los cabeceros están bien puestos y contienen correctas descripciones de lo que se encuentra en la página.

La densidad de términos no es muy elevada 3% aunque se encuentra en la media de las páginas que se han visitado anteriormente. No obstante la Url y los títulos son descriptivos ya que el nombre de la marca ya es un referente a nivel nacional sobre la actividad de la página. No obstante faltan algunas metadescripciones en las imágenes.

El "PAGE AUTHORITY" de esta página es alto (35,64) ya que está por encima de la media de las páginas que se han explorado anteriormente. Su valor en el rango Alexa es 932.262 por lo que está bien posicionada.

2.6.5.1 Términos clave

La densidad de los términos repetidos llega al 3% lo cual no es mucho, aunque suele ser la media comparada a las demás páginas consultadas. Los términos clave que se han encontrado en estas páginas son: Osborne, bebidas, bebidas espirituosas, alimentación, ibéricos, jamón ibérico, bellota, bodegas, vinos de jerez, vinos de rioja, oportos y cinco jotas jamón.

2.6.5.2 Redes sociales

La presencia de esta página web en las redes sociales es notable ya que está en la mayoría de las redes sociales más importantes como Facebook, Google+ y LinkedIn. En Facebook es particularmente activa ya que tiene 170 “me gusta” y ha sido compartida 65 veces.

Destaca también una fuerte presencia en LinkedIn, ésto le ayuda mucho ya que esta red social está bien valorada por Google. Por otro lado la participación en Google+ es muy baja.

2.6.5.3 Enlaces

Esta página posee un número significativo de enlaces, ya que a día 30/03/2016 posee un total de 1076 backlinks que le ayudan a escalar en Google puesto que algunos de estos enlaces son de páginas oficiales con un “PAGE AUTHORITY” de más del 30%.

Además cuenta con 41 enlaces internos, una cifra bastante alta y otros 14 enlaces externos. Todo ello le ayuda a conseguir un mejor posicionamiento.

Si se analizan minuciosamente los “backlinks” de la página se pueden encontrar enlaces con un “PAGE AUTHORITY” muy alto. Entre ellos está la página de la fundación Osborne que tiene un “PAGE AUTHORITY” del 58,76.

También se referencian en otras empresas que ofrecen servicios de posicionamiento y herramientas SEO. Otras páginas que clasifican páginas por conceptos, foros y periódicos. Estos enlaces son de alta calidad lo que le van a asegurar un buen posicionamiento.

2.6.6 Tudespensa (<https://www.tudespensa.com/>)

Tudespensa.com es una página que se dedica a vender un amplio abanico de productos. Entre ellos se pueden dividir los productos típicos que se encuentran en un supermercado (aceite, lácteos, panadería, etc.), productos frescos (pescadería, carnicería y frutería), productos ecológicos, utensilios de cocina y parafarmacia.

Esta página cumple con casi todas las recomendaciones SEO, ya que contiene títulos descriptivos y los encabezamientos están correctamente escritos con términos descriptivos. Las meta descripciones están puestas, no obstante no se han identificado algunas de las imágenes.

La densidad de términos no es muy alta ya que los términos más descriptivos apenas llegan al 1,3%. Por lo que se puede deducir que esta página se enfoca más en el entorno puramente visual y no tanto en el texto. No obstante, el título de la página es muy descriptivo y ayuda a entender el objetivo de la página.

Esta página posee un “PAGE AUTHORITY” significativo (39,98), el rango de Alexa es significativo también, ya que es 153.782 lo que supone uno de los mayores que se han visto a lo largo de este estudio. Estos dos elementos ayudan enormemente al buen posicionamiento de la página en Google.

2.6.6.1 Términos clave

La densidad de los términos clave no es muy alta ya que no se supera el 1,3%. Por lo que no se ha hecho mucho hincapié en colocar muchos términos clave en el texto. Las palabras clave que se han encontrado son: supermercado online, super online, compra online, supermercado internet.

2.6.6.2 Redes sociales

Esta página cuenta con una presencia modesta en las redes sociales. Está presente en la mayoría de ellas como Facebook, Google+, Pinterest y LinkedIn. Esta página destaca en LinkedIn donde tiene mucho peso, en menor medida también en Facebook con varios “Me gusta” y compartiendo varios enlaces.

2.6.6.3 Enlaces

El número de enlaces internos de esta página es elevado ya que a fecha de hoy (2/4/16) cuenta con 56 enlaces internos y 12 externos. Cuenta además con 8.691 “backlinks” lo que le aporta una ventaja muy significativa en su posicionamiento.

Algunos de estos enlaces tienen un “PAGE AUTHORITY” de 46,23%. Lo que los convierte en páginas reconocidas dentro de Google. Algunos de las páginas que le apuntan son Bancos, ONG, foros, periódicos, blogs, etc. Muchas de ellas muy activas y bastante populares a nivel nacional.

2.6.7 Estabuenisimo (<http://www.estabuenisimo.es/>)

La página “estabuenisimo.com” se concentra en vender productos de alta calidad abarcando un amplio abanico de productos. Que van desde jamones, embutidos, conservas, foie, patés, salsas, cremas, sal, especias, legumbres hasta vinos y licores.

También ofrece “cestas” que no son otra cosa que paquetes con la combinación de varios productos con precios más atractivos. La presentación de la página es sencilla pero clara y facilita la información de contacto.

Esta página cumple con la mayoría de las recomendaciones SEO, ya que la Url tiene una longitud correcta, los títulos y los cabeceros son descriptivos. Las metadescripciones son correctas, aunque las imágenes no están descritas correctamente.

La densidad de los términos no es demasiado alta ya que solo llega hasta 2,42% el término clave más repetido. El resto de los términos apenas supera el 1%, la página tiene un entorno muy gráfico y deja un espacio muy ajustado para las descripciones.

El “PAGE AUTHORITY” de la página es significativamente alto (34,58), el “Alexa Rank” es 5.619.038 que es significativo también ya que en comparación a otras páginas está en una franja que supera la media de las otras páginas.

2.6.7.1 Términos clave

La densidad de la repetición de los términos clave no es muy alta ya que apenas se logra superar el 2,42% en un par de términos clave. Esto es debido a que la página tiene un entorno muy sencillo con grandes imágenes para dar más importancia al aspecto visual que al descriptivo.

La familia de términos más destacados que se ha extraído es: regalos gourmet, comprar productos gourmet, tienda gourmet, productos delicatessen, venta de jamón, venta de vino online, venta de queso, venta de aceite, cestas de navidad, productos.

2.6.7.2 Redes sociales

Esta página cuenta con una fuerte presencia en Facebook y en menor medida en Google+. En Facebook tiene 38 “me gusta” y varios comentarios. Su página ha sido compartida 42 veces lo que le ayuda a posicionarse en Google.

Se puede concluir que aunque tiene presencia en las redes sociales solo se ha concentrado en una ya que en Google + tiene poca presencia al estar compartida su página solo 5 veces.

2.6.7.3 Enlaces

Esta página cuenta con un gran número de enlaces internos (15) y otros muchos externos (12). No obstante la verdadera diferencia la marcan los 603 “backlinks” que tiene la página, lo que le ayuda a escalar en su posicionamiento en Google.

La página cuenta con enlaces desde distintas páginas como son foros sobre recetas, otras tiendas online de productos de distinta clase, un portal para facilitar la información de empresas.

2.6.8 Singluten (<http://singluten.com/>)

La página “singluten.com” es un supermercado en línea que ofrece una amplia gama de productos que se encuentran en cualquier supermercado. Ya que abarca desde fruta, menaje del hogar, libros, postres, chocolates, etc.

La página cumple la mayoría de las recomendaciones SEO básicas como son añadir los metadatos en cada página, colocar cabeceros de las páginas, además los títulos son descriptivos de lo que se está anunciando en cada página. No obstante no ha incluido la etiqueta “alt” en las imágenes para dar información acerca de su contenido.

La densidad de términos es alta ya que ronda el 2,2% de densidad de términos, aunque algunas palabras como “sin gluten” destacan, ya que consiguen un 4,25%. En cuanto al rango Alexa “Alexa Rank” es 1.390.947, el cual es relativamente alto en comparación a las otras páginas estudiadas.

El “PAGE AUTHORITY” de la página no obstante es alto 41,37. Es uno de los valores más altos que se han encontrado si se compara con las anteriores páginas estudiadas. Por lo tanto el “PAGE AUTHORITY” y el rango Alexa ayudan positivamente a la página a conseguir un buen posicionamiento.

2.6.8.1 Términos clave

La densidad de los términos es alta comparado a las otras páginas estudiadas ya que algunos términos clave como gluten consiguen un 4,25%, otro caso es “arroz” con 1,52% y “harina” con 2,26%. Esta página se apoya mucho en lo visual, no obstante se han escogido bien los términos para posicionar cada una de las distintas secciones.

La familia de términos que se han recogido son: gluten, alimentos sin gluten, levadura, celíacos, tienda online, galletas, arroz, sin gluten y harina.

2.6.8.2 Redes sociales

La presencia de esta página en las redes sociales es muy activa. Está presente en todas las redes sociales principales. Tiene mucha fuerza en Facebook donde acumula centenares de comentarios y “me gusta”.

También tiene presencia en Google+, Pinterest y LinkedIn lo que la hace una página reconocida por todos estos canales gratuitos. Lo que le da mucha relevancia y ayuda a su posicionamiento en Google.

2.6.8.3 Enlaces

Esta página cuenta con un gran número de enlaces internos ya que cuenta con 60 enlaces, también cuenta solo con 3 enlaces externos. No obstante esta página destaca por los 24.383 “backlinks” que tiene, ya que es el valor más alto que se ha registrado de todas las páginas estudiadas.

Los enlaces que tiene la página son de alta calidad ya que cuenta con enlaces de blogs de comercio electrónico, algunos periódicos online, otras tiendas online y la propia página de ecommerce-news.

2.7 Marco regulador

La política de uso de Facebook permite explorar solamente aquellas páginas de empresas comerciales de forma gratuita. Para poder explorar los perfiles privados o públicos de los usuarios es necesario comprar una API que solamente posee Facebook.

No obstante también se publican en la página oficial de desarrolladores una serie de directrices y normas para los programadores y usuarios de las APIs que Facebook ofrece. Se detallan a continuación de forma resumida los principales aspectos de esta [80].

- Crear un producto de calidad: en el cual la aplicación sea estable, que siga las políticas de privacidad, las políticas de publicidad, mantener actualizadas la descripción y la categoría de la aplicación y no defraudar.
- No enviar a los usuarios mensajes desde la aplicación de Facebook.
- Proteger los datos de que son accedidos desde la aplicación.
- Uso correcto de las funciones de Facebook, que ayuden a enriquecer el contenido de la página y facilitar la promoción de Facebook.
- Cumplir las leyes de protección de datos y derechos de autor de España y otros países donde se apliquen estas [81].

3 Análisis de la solución

En este capítulo se procede a definir de forma precisa lo que la aplicación debe desarrollar. Se estudiarán con detalle las necesidades que se van a satisfacer. Para ello se debe analizar el sector de la alimentación en internet para poder comprender de una forma más sencilla el comportamiento de éste y las claves de su operativa.

Se analizarán los diferentes términos o microlenguajes de este campo para poder hacer un análisis de los términos variables más significativos que usan las páginas líderes del mercado.

El campo de la investigación se centrará en Facebook y en otras páginas externas a éste que han sido catalogadas por la asociación de clubecommerce en la siguiente fuente [10]. No solo se ha consultado esta fuente sino que se ha contrastado otra más para operar con la mayor diversidad de resultados, consultando esta otra fuente [11].

Se ofrece un estudio del mercado de las aplicaciones SEO genéricas que hay en el mercado y la viabilidad que hay en aplicaciones de este estilo. Se propone estudiar la frecuencia de uso de una serie de términos clave a los cuales se referirán como parte de un diccionario de palabras relevantes.

Este diccionario contiene una selección de palabras que han sido extraídas de las principales páginas web del mercado. Estas páginas pueden ofrecer pistas de como han conseguido un posicionamiento tan elevado.

3.2 Definición del sistema

3.2.1.1 Determinación del alcance del sistema

Tanto el sistema construido como el presente documento pretenden crear una herramienta capaz de rastrear los principales términos usados en las páginas web del mundo de la alimentación en España.

El propósito es rastrear las principales páginas abiertas de Facebook para conocer la frecuencia de repetición de estos términos relevantes. El conocimiento de uso de estos términos a lo largo del tiempo ayudará al usuario a poder dirigir sus campañas de publicidad y crear su sitio web.

La elección de los términos clave para una página no es una decisión clara y no atiende a unos criterios estrictamente precisos por lo que el sistema planteado pretende servir como una herramienta de apoyo para que el usuario encuentre estos términos.

Esta decisión de emplear unos términos u otros influye directamente en el impacto económico y de posicionamiento de la página. Por todo esto se considera un factor clave el consolidar bien estos términos y que tengan relación con el tema de la página.

Por ello se ha creado un rastreador de la red social Facebook, que busca estos términos en unas páginas predeterminadas por el desarrollador para conseguir obtener las frecuencias de uso y mostrarlas en un gráfico.

El vocabulario de las redes sociales no está limitado, es por ello que no basta con lanzar un simple análisis por la red, sino que hay que determinar sí el comentario que se está analizando es relevante para ser analizado o no. Es por ello que se emplea una medida llamada TF para analizar la relevancia del comentario.

El algoritmo TF (Term Frequency) mide de una forma numérica como de relevante es una palabra para un determinado documento o en este caso un comentario. En el sistema se usa como factor para medir la aparición de los términos clave, es decir, la ponderación que tiene un término concreto en un determinado comentario.

Esta ponderación es muy importante ya que en función de esta, se determina si un comentario es relevante o no para ser analizado después para llevar a cabo el análisis de frecuencias que se mostrará al cliente.

Este valor de TF aumenta cada vez que un determinado término clave aparece en el texto o comentario, no obstante al tratarse de un coeficiente entre las veces que aparece el término clave en el comentario y el número de palabras del comentario.

Otro algoritmo usado para el desarrollo del sistema ha sido el algoritmo PART, este algoritmo es un clasificador que el desarrollador ha utilizado para comprobar la efectividad del análisis automático realizado por el sistema.

Este análisis de efectividad del sistema es vital ya que los datos que se muestran al usuario dependen directamente del control que ejerce el algoritmo TF sobre el análisis de los comentarios.

Si la efectividad de ese análisis sobre los comentarios usando el algoritmo TF fuese pobre, los datos que se presentan al cliente carecerían de validez real ya que se estarían examinando comentarios que puede que no fuesen relevantes para la cuestión a tratar. La cuestión a tratar en este caso es la repetición de términos del sector de la alimentación.

El algoritmo PART se emplea dentro del entorno de WEKA para análisis de minería, este entorno reside fuera del subsistema que se ha creado. Ya que WEKA es un sistema en si que se usa para analizar la efectividad de los análisis que ejecuta el sistema.

El análisis de efectividad se lleva a cabo rellenando un documento con un formato especial de WEKA conocido como “.ARFF”. El cual permite a WEKA analizar la información contenida en los análisis ejecutados de forma automática a partir de unas reglas que ha creado previamente.

Las reglas que usa el algoritmo PART se generan a partir de un modelo predefinido por el usuario. Este modelo se genera con unas clases de entrenamiento que el desarrollador analiza manualmente leyendo los comentarios y juzgando si estos son o no relevantes para ser analizados.

Una vez se juzga si los comentarios son relevantes o no, se introduce toda la información relativa a ellos dentro de un documento con el formato especial “.ARFF”. Este documento es analizado por WEKA para extraer las reglas del modelo analizado.

Las reglas del modelo analizado servirán para medir después la efectividad del módulo de análisis automático que lleva a cabo el sistema de todos los comentarios que recoge de las páginas de la red social Facebook.

3.2.1.2 Identificación del entorno tecnológico

El sistema desarrollado consiste en un servicio web para rastrear comentarios en las páginas de la red social de Facebook y mostrar la frecuencia de uso de unos determinados términos. Por lo que los requisitos de la máquina del cliente deben coincidir con los del entorno donde se desarrolló dicha aplicación.

- Java JDK versión 8 instalado.
- Memoria RAM de 1 GB dedicada, se recomienda que el equipo tenga por lo menos 4GB
- Procesador a 2.2 GHz mínimo.

- 128 GB de disco duro reservado
- Sistema operativo Windows 10 (recomendado).

3.2.1.3 Especificación de estándares y normas

A continuación se presentan las restricciones generales que se han tenido en cuenta para realizar el proyecto, las dependencias existentes y el entorno operacional que se requiere para hacer funcionar el sistema.

Restricciones generales

Se detallan las siguientes restricciones:

- Todas las clases que ejecutan instrucciones sobre la base de datos deben heredar de la librería (“java.sql.*;”).
- Todas las clases que interactúan con la API de Facebook deben heredar las librerías (“com.restfb.*; ”), y sus respectivas variantes para poder ejecutar los token de acceso y las instrucciones pertinentes.
- Todas las clases que interactúan con la aplicación de JfreeChart para llevar a cabo las tareas de diagramado y presentación deben incluir la librería (“org.jfree.chart.*; ”) y sus respectivas variantes a la vez que la librería (java.awt.*;).
- Se aconseja escribir comentarios explicativos para las distintas clases y métodos para facilitar la comprensión y crear una documentación clara.

Dependencias y supuestos

El sistema que se ha desarrollado no es una herramienta SEO completa. No obstante los módulos de navegación son suficientes como para permitir al usuario interactuar con la aplicación de una forma clara y rápida.

Se ha revisado la documentación del sistema XAMPP, que es la base sobre la que se ejecuta el sistema para seguir la guía de buenas prácticas que se encuentran en el siguiente enlace (<https://www.apachefriends.org/community.html>) [13].

Para establecer la exploración dentro de Facebook se usa una API ofrecida por la página de desarrolladores de Facebook. Para ello el desarrollador se ha registrado en la página con un perfil de empresa y ha registrado la API siguiendo la información sugerida por Facebook en el siguiente enlace (<https://spring.io/guides/gs/register-facebook-app/>).

Se solicitará un número de registro o token dentro de esta página que se utilizará como una llave de acceso dentro de la página. Debe aparecer algo muy parecido a lo que se muestra a continuación.

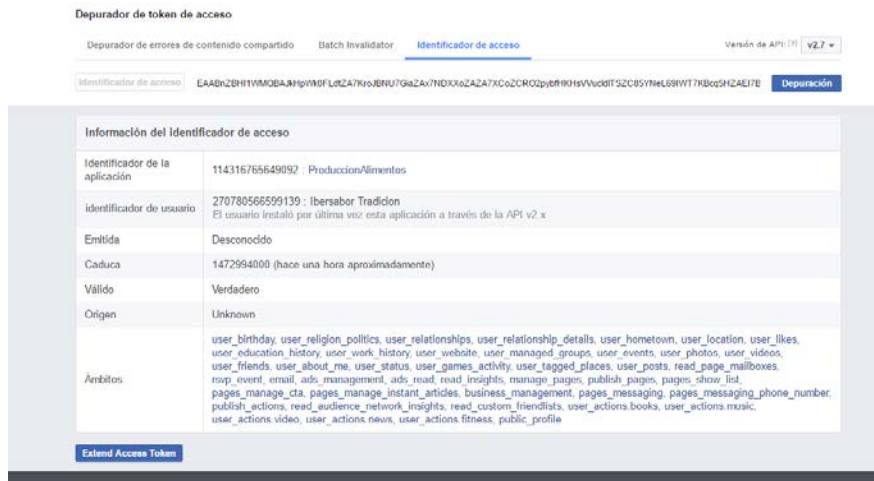


Ilustración 8. API Facebook. Fuente [79].

Entorno operacional

Cabe mencionar que los equipos necesarios para desarrollar el código de la aplicación han tenido siempre la misma configuración por lo que se recomienda seguir siempre la misma. Se detalla a continuación la configuración que se ha usado:

- Windows 10 64 bits
- XAMPP versión 5.6.20 (engloba a Netbeans y MariaDB)
- Microsoft Office 2010
- Weka (cualquier versión)
- API de Facebook versión 2.0 .

3.2.1.4 Usuarios participantes (stakeholders) y usuarios finales.

A continuación se detallan los dos tipos de usuarios que son grupo de interés para el desarrollo de la aplicación. El primer grupo de interesados serían los Stakeholders que son aquellos interesados en que se lleve a cabo el desarrollo de la aplicación. En este caso correspondería a la UC3M ya que es la universidad donde se realiza el TFG y donde se va a presentar.

El otro grupo de usuarios correspondería a los usuarios finales de la aplicación gratuita para rastrear Facebook que son los que realmente harán uso del sistema.

3.3 Requisitos

Esta sección tiene como objetivo exponer los requisitos de la solución final. Para detallar dichos requisitos se emplearán una serie de tablas comunes para todos ellos. El desarrollador se ha basado en su experiencia previa y ha consultado a su tutor para acordar los más importantes. La tabla que se muestra se ha usado por el desarrollador durante la carrera, por lo que se reutiliza.

La clasificación de los requisitos se hará en función de si éstos son funcionales o no funcionales. De los requisitos no funcionales se analizarán requisitos de rendimiento, interfaz, disponibilidad, compatibilidad, portabilidad, mantenibilidad, usabilidad y comunicación

. La tabla que se va a usar se muestra a continuación con la debida explicación de cada campo.

- **Identificador:** Campo alfanumérico que identifica a un requisito de forma única. Está formado por una etiqueta que determina el tipo de requisito del que se trata y un número. Podemos elegir entre las siguientes etiquetas para especificar el tipo de requisito:
 - **RF:** Se corresponde con el acrónimo de Requisito Funcional.
 - **RNF:** Se corresponde con el acrónimo de Requisito No Funcional.
- **Nombre:** Campo que describe de manera resumida el requisito.
- **Prioridad:** se detalla el grado de relevancia del requisito, en una escala de baja, media y alta prioridad.
- **Necesidad:** se valora el impacto de no incluir el requisito en el sistema, diferenciando entre esencial, deseable y opcional.
- **Estabilidad:** este campo describe la duración del requisito respecto la vida del software.
- **Descripción:** campo que detalla específicamente las características de cada requisito.

3.3.1.1 Requisitos funcionales

Los requisitos funcionales son los que describen la funcionalidad que se requiere para el sistema. Se redactan en lenguaje natural, para así evitar en la medida de lo posible la ambigüedad o la imprecisión al redactarlos.

Identificador	RF - 01
Nombre	Carga de datos de comentarios(posts) de Facebook inicial en la base de datos
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	Se deben cargar los comentarios (posts) de las páginas de Facebook seleccionadas del año 2015.

Tabla 2- Carga posts inicial

Identificador	RF - 02
Nombre	Carga de datos de palabras del diccionario de palabras relevantes en el campo de la alimentación.
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	Se deben cargar las palabras relevantes para poder efectuar los análisis de relevancia oportunos.

Tabla 3- Carga diccionario palabras relevantes

Identificador	RF - 03
Nombre	Carga de datos de palabras negativas
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	Se deben cargar las palabras de carácter negativo para que se pueda valorar la relevancia del comentario.

Tabla 4- Carga de datos de palabras negativas

Identificador	RF - 04
Nombre	Mostrar tabla de frecuencias
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe ser capaz de mostrar en una grafica las frecuencias de uso de la palabra introducida por el usuario.

Tabla 5- Mostrar tabla de frecuencias

Identificador	RF - 05
Nombre	Usuario debe introducir palabra deseada
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe ser capaz de permitir al usuario introducir una palabra del catálogo de palabras relevantes para que pueda ejecutar la búsqueda de la frecuencia de uso.

Tabla 6- Usuario debe introducir palabra deseada

Identificador	RF - 06
Nombre	Mostrar palabras relevantes categoría Alimentación
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe mostrar al usuario las palabras relevantes de la categoría general de Alimentación.

Tabla 7- Mostrar palabras relevantes categoría Alimentación

Identificador	RF - 07
Nombre	Mostrar palabras relevantes categoría Bebidas
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe mostrar al usuario las palabras relevantes de la categoría general de Bebidas.

Tabla 8- Mostrar palabras relevantes categoría Bebidas

Identificador	RF - 08
Nombre	Mostrar palabras relevantes categoría Regiones
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe mostrar al usuario las palabras relevantes de la categoría general de Regiones.

Tabla 9- Mostrar palabras relevantes categoría Regiones

Identificador	RF - 09
Nombre	Mostrar palabras relevantes categoría Términos especiales
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe mostrar al usuario las palabras relevantes de la categoría general de Términos especiales.

Tabla 10- Mostrar palabras relevantes categoría Términos especiales

Identificador	RF - 10
Nombre	Mostrar palabras relevantes categoría Comidas diarias
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe mostrar al usuario las palabras relevantes de la categoría general de comidas diarias.

Tabla 11- Mostrar palabras relevantes categoría Regiones

Identificador	RF - 11
Nombre	Determinar la relevancia de un comentario (post)
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe determinar la relevancia o no de un comentario para su posterior análisis de frecuencia.

Tabla 12- Determinar la relevancia de un comentario (post)

Identificador	RF - 12
Nombre	Analizar frecuencia de post relevante
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe analizar la frecuencia de aparición de un término seleccionado por el usuario que aparezca en los comentarios (posts) que hayan sido catalogados como relevantes.

Tabla 13- Analizar frecuencia de post relevante

Identificador	RF - 13
Nombre	Mostrar página de índice
Prioridad	Alta prioridad
Necesidad	Esencial
Estabilidad	Alta(largo)
Descripción	El sistema debe mostrar la página del índice cuando el usuario lo requiera.

Tabla 14- Mostrar página de índice

Identificador	RF – 14 (opcional)
Nombre	Usuario puede aconsejar páginas de Facebook a analizar.
Prioridad	Baja
Necesidad	Opcional
Estabilidad	Media
Descripción	El usuario debe poder aconsejar al gestor de mantenimiento del sistema, nuevas páginas de Facebook que analizar

Tabla 15- Usuario puede aconsejar páginas de Facebook a analizar.

3.3.1.2 Requisitos no funcionales

Los requisitos no funcionales son aquellos que describen las restricciones o criterios que pueden usarse para analizar la operación del sistema a la hora de realizar el software. Las restricciones abarcan varios tipos que se describen a continuación: [20]

- Usabilidad (Usability)
- Portabilidad (Portability)
- Interfaz (Interface)
- Comunicación (Communication)
- Rendimiento (Performance)
- Disponibilidad (Availability)

- Hardware
- Software

SOFTWARE

Identificador	RNF – S – 01
Nombre	Sistema operativo Windows
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	El sistema operativo aconsejado para hacer funcionar este sistema es Windows 10.

Tabla 16- Sistema operativo Windows

HARDWARE

Identificador	RNF – H – 01
Nombre	Procesador recomendado
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	Se recomienda utilizar un procesador como mínimo de 5 núcleos a 2.2GHz

Tabla 17 - Procesador recomendado

Identificador	RNF – H – 02
Nombre	Memoria RAM recomendada
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	Se recomienda utilizar una memoria RAM de 4 GB

Tabla 18- Memoria RAM recomendada

RENDIMIENTO

Identificador	RNF – PE – 01
Nombre	El tiempo de respuesta
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	El tiempo de respuesta del sistema ante una consulta del usuario será menor a 10 minutos

Tabla 19- El tiempo de respuesta

DISPONIBILIDAD

Identificador	RNF – D – 01
Nombre	Disponibilidad anual
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	La disponibilidad anual será de al menos 340 días al año.

Tabla 20- Disponibilidad anual

USABILIDAD

Identificador	RNF – U – 01
Nombre	Claridad de diseño
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	Deben quedar claramente separadas las distintas opciones del sistema mediante enlaces y botones si se cree conveniente.

Tabla 21- Claridad de diseño

PORTABILIDAD

Identificador	RNF – P – 01
Nombre	Portabilidad del sistema entre sistemas operativos
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	El sistema debe funcionar en otros equipos que tengan instalado desde la versión Windows 10 en adelante.

Tabla 22- Portabilidad del sistema entre sistemas operativos

Identificador	RNF – P – 02
Nombre	Interacción entre los distintos subsistemas
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	El sistema XAMPP debe interactuar con el sistema de ficheros de Windows para guardar los comentarios (posts).

Tabla 23- Interacción entre los distintos subsistemas

INTERFAZ

Identificador	RNF – I – 01
Nombre	El diseño de la interfaz consistente
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	El diseño de la interfaz será consistente con las interfaces del resto de subsistemas.

Tabla 24- Interacción entre los distintos subsistemas

Identificador	RNF – I – 02
Nombre	Pantalla inicial
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	La pantalla inicial está compuesta por varios enlaces a las distintas páginas, las cuales muestran los distintos términos con los que se puede trabajar.

Tabla 25- Pantalla inicial

Identificador	RNF – I – 03
Nombre	Pantalla con gráfica de resultados
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	Se debe ejecutar una prueba aleatoria que debe mostrar los resultados de la palabra buscada en una gráfica en el centro de la página.

Tabla 26- Pantalla con gráfica de resultados

Identificador	RNF – I – 04
Nombre	Pantalla con términos de Bebidas
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	En la página de Bebidas se deben mostrar todos los términos clave de las bebidas que se ofrecen en el sistema.

Tabla 27- Pantalla con términos de Bebidas

Identificador	RNF – I – 05
Nombre	Pantalla con términos de Alimentos
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	En la página de Alimentos se deben mostrar todos los términos clave de los alimentos que se ofrecen en el sistema.

Tabla 28- Pantalla con términos de Alimentos

Identificador	RNF – I – 06
Nombre	Pantalla con términos de Regiones de España
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	En la página de Regiones de España se deben mostrar todos los términos clave de las regiones de España que se ofrecen en el sistema.

Tabla 29 - Pantalla con términos de Regiones de España

Identificador	RNF – I – 07
Nombre	Pantalla con términos de Términos especiales
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	En la página de Términos especiales se deben mostrar todos los términos clave de los términos especiales que se ofrecen en el sistema.

Tabla 30- Pantalla con términos de Bebidas

Identificador	RNF – I – 08
Nombre	Pantalla con términos de Comidas diarias
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	En la página de Comidas diarias se deben mostrar todos los términos clave de las comidas diarias que se ofrecen en el sistema.

Tabla 31- Pantalla con términos de Comidas diarias

Identificador	RNF – I – 09
Nombre	Interfaz sencilla e intuitiva
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	La interfaz será sencilla e intuitiva para que el usuario saque el mayor rendimiento a través de ella.

Tabla 32- Pantalla con términos de Comidas diarias

Identificador	RNF – I – 10
Nombre	Notificación no encontrado
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	Se le notificará al usuario si la palabra que ha escrito para lanzar la consulta no se ha encontrado en la base de datos.

Tabla 33- Pantalla con términos de Comidas diarias

COMPATIBILIDAD

Identificador	RNF – C – 01
Nombre	Comunicación entre subsistemas
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	El Interfaz gráfica se comunica con la base de datos a través de un Servlet el cual a su vez comunica con un pool de conexiones con la base de datos.

Tabla 34- Comunicación entre subsistemas

Identificador	RNF – C – 02
Nombre	Compatibilidad con otras aplicaciones
Prioridad	Alta
Necesidad	Esencial
Estabilidad	Alta (largo)
Descripción	El sistema desarrollado es compatible con otros sistemas como el navegador o el ordenador anfitrión (Host).

Tabla 35- Compatibilidad con otras aplicaciones

3.3.1.3 Especificación de casos de uso

En el siguiente apartado se procede a examinar los casos de uso, para ello se utilizarán ilustraciones para facilitar la comprensión de éstos, junto a tablas para detallar más la descripción, como los escenarios donde se utilizarán.

Los casos de uso son textos narrativos que describen el comportamiento de un sistema desde el punto de vista de un actor, persona o sistema que usa el sistema para conseguir un determinado

propósito. En el supuesto caso de que el actor no pueda completar el objetivo concreto que plantea, entonces no se podrá representar ese caso de uso.

Se presentan dos apartados para mostrar de forma detallada y concisa los casos de uso. En el primer sub apartado se muestran los diagramas de uso con un diagrama UML. En el segundo apartado se muestran los casos de uso en tablas de forma más detallada.

Diagrama de casos de uso

Se presentan a continuación los casos de uso del sistema así como los actores que interactúan con él. Para su elaboración se ha seguido la notación UML (Unified Modeling Language), la cual queda explicada a continuación.

Los círculos corresponden a los casos de uso y en su interior se encuentra el nombre que se le ha asignado a cada uno de ellos. Los actores son representados con una figura humana con su rol escrito encima. La unión de un determinado actor con un caso de uso se representa con una línea.

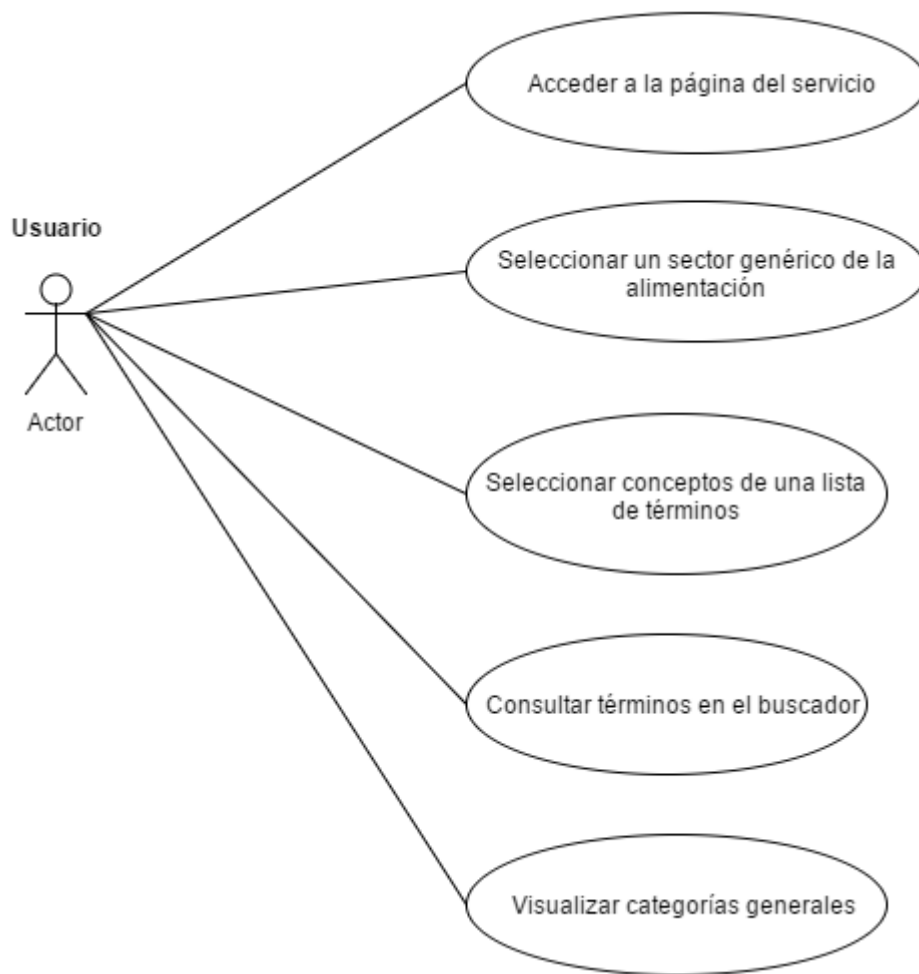


Ilustración 9. Casos de uso. Elaboración propia

Tablas de casos de uso

En este subapartado se presentan los casos de uso identificados mediante una tabla con un formato más extendido. Se continuará empleando la notación UML, indicando para cada uno de los casos de uso los siguientes campos.

- **Identificador:** Código alfanumérico que define de forma unívoca un caso de uso. Seguirá la nomenclatura CDU-XX, donde CDU se refiere a las iniciales de Caso De Uso y XX al número que corresponda.
- **Nombre de caso de uso:** Calificativo unívoco del caso de uso.
- **Actor/es que participa en el caso de uso:** Personas o sistemas externos al desarrollado en este proyecto que interactúan con el sistema para conseguir un propósito.
- **Tipo de caso de uso:** la clasificación del caso de uso según su tipo se hará mediante dos criterios:
 - **Según la importancia del caso de uso en el proyecto:** el caso de uso podrá ser primario, secundario u opcional.
 - **Según el grado de compromiso con el código o implementación:** el caso de uso puede ser esencial o real. Es esencial cuando en él se

especifiquen requisitos tecnológicos o se muestran detalles de su codificación.

- **Descripción:** explicación del comportamiento que representa el caso de uso.
- **Pre-condiciones:** conjunto de condiciones que deben ser ciertas antes de iniciar el caso de uso: Es muy común que la pre-condición sea el fruto de un caso de uso anterior, en estos casos se pondrá como pre-condición el caso de uso que ha debido darse con anterioridad al que está siendo descrito.
- **Post-condiciones:** indican el estado final del sistema después de que el caso de uso termine con éxito.
- **Propósito:** intención del caso de uso.

C-01	Acceder a la página del servicio
Actor/es	Usuario.
Tipo	Primario y esencial.
Descripción	El usuario debe poder acceder al servicio proporcionado por el sistema a través de su navegador para utilizar sus servicios.
Pre-condiciones	Conexión a Internet, el servicio debe estar en línea.
Post- condiciones	El usuario debe poder visualizar el índice.
Propósito	El usuario debe poder acceder al servicio ofrecido por el sistema para poder visualizar las distintas opciones que se le ofrecen.

Tabla 36- C01

C-02	Seleccionar un sector genérico de la alimentación.
Actor/es	Usuario.
Tipo	Primario y esencial.
Descripción	El usuario debe poder seleccionar una de las clases genéricas que se ofrecen en el índice.
Pre-condiciones	Conexión a Internet, el servicio debe estar en línea y el usuario debe haber accedido al índice.
Post- condiciones	El usuario debe poder acceder a la clase genérica seleccionada.
Propósito	El usuario debe poder acceder a la clase genérica que ha seleccionado para ver los términos clave que se ofrecen dentro de esa clase.

Tabla 37- C02

C-03	Seleccionar conceptos de una lista de términos
Actor/es	Usuario.
Tipo	Primario y esencial.
Descripción	El usuario debe poder seleccionar los conceptos de una lista de términos la cual podrá explorar desplazándose por la página donde estos se encuentran.
Pre-condiciones	Conexión a Internet, el servicio debe estar en línea y la página de la categoría debe estar cargada en el navegador.
Post- condiciones	Página de la categoría debe estar cargada incluso después de hacer la lectura.
Propósito	El usuario debe poder leer todos los términos que se le ofertan con fines informativos. Uno de estos términos será seleccionado por el usuario para realizar su búsqueda. Este proceso puede ser repetido todas las veces que el usuario quiera.

Tabla 38-C03

C-04	Consultar términos en el buscador
Actor/es	Usuario.
Tipo	Primario y esencial.
Descripción	El usuario debe poder escribir la palabra clave
Pre-condiciones	Conexión a Internet, el servicio debe estar en línea y haber accedido a la página de la clase genérica que seleccionó previamente.
Post- condiciones	El usuario debe poder escribir el término y enviar la petición al sistema.
Propósito	El usuario debe poder introducir el término clave deseado para que el sistema le devuelva los resultados de esa palabra que se ha introducido.

Tabla 39- C04

C-05	Visualizar categorías generales
Actor/es	Usuario.
Tipo	Primario y esencial.
Descripción	El usuario debe poder visualizar las categorías generales cuando desee hacerlo.
Pre-condiciones	Conexión a Internet, el servicio debe estar en línea.
Post- condiciones	Volver a la página del índice.
Propósito	El usuario debe poder visualizar las categorías generales que se encuentran en el índice, para ello se le ofrece un enlace dentro de la página donde esté para que pueda visualizarlas.

Tabla 40- C05

3.3.1.4 Matriz de trazabilidad

La matriz de trazabilidad permite comprobar que todos los requisitos han sido extraídos a partir de las historias de usuario. Se presenta una tabla a continuación:

	C-01	C-02	C-03	C-04	C-05
RF- 01	x			x	
RF -02	x		x	x	
RF- 03	x				
RF- 04				x	
RF-05				x	
RF- 06		x			
RF- 07		x			
RF- 08		x			
RF- 09		x			
RF- 10		x			
RF- 11				x	
RF- 12				x	

Tabla 41- Matriz de trazabilidad

3.4 Identificación de subsistemas de análisis

El sistema que se ha construido es bastante complejo por lo que para repartir las tareas que debe hacer se ha repartido esta complejidad entre varios subsistemas. Las funcionalidades se han repartido convenientemente en función de las librerías que se deben usar para implementar su código y por la tarea que desempeñan.

Los subsistemas que se han propuesto son:

- **Interfaz gráfica:** tiene la función de mostrar la interfaz de usuario a todos los demás subsistemas.
- **Motor de análisis:** es el encargado de representar toda la funcionalidad de la aplicación que opera con el analizador de términos. Este sistema es muy complejo y grande por lo que se ha dividido en subsistemas:
 - **MariaDB:** se encarga de guardar todos los datos y asegurar la persistencia de las tablas de datos con los comentarios (posts), las palabras relevantes y las negativas.
 - **Analizador de comentarios:** este módulo se encarga de analizar los comentarios guardados en la base de datos, clasificándolos en función de si son relevantes o no.
 - **Convertor a tipo Weka:** se encarga de convertir la información del analizador a un texto interpretable por Weka para medir la tasa de éxito del análisis.
 - **Palabras positivas y negativas:** se encarga de cargar el diccionario de palabras positivas y el de las negativas para que se pueda hacer el análisis de relevancia.

- **Renderizador de gráficas:** se encarga de crear las gráficas donde se imprimen los datos que se han extraído del análisis de comentarios.

Se presenta a continuación el diagrama de componentes.

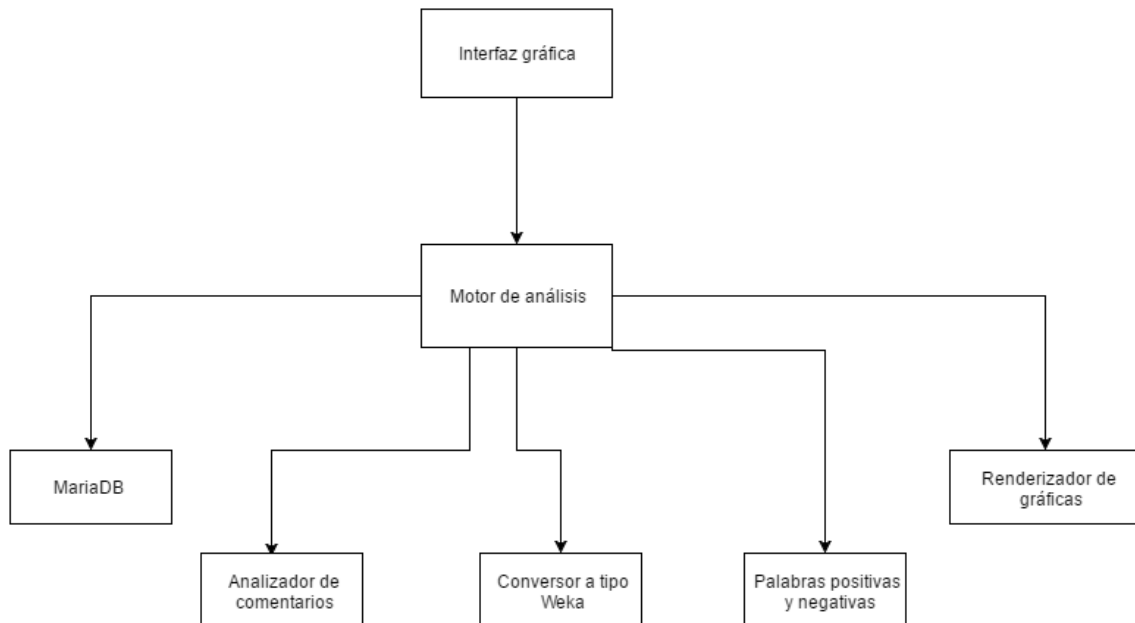


Ilustración 10. Diagrama de componentes

4 Diseño

4.1 Arquitectura del sistema

A continuación se analiza la arquitectura del sistema, para ello se va a llevar a cabo un diagnóstico del módulo o módulos desarrollados con el resto de sistemas con los que convive. Se establecerá una norma para el desarrollo del código.

4.1.1.1 Diseño de los niveles de arquitectura

El sistema que se ha desarrollado sigue una arquitectura de 3 capas, esta arquitectura está diseñada para solventar los problemas de escalabilidad que suelen tener otras arquitecturas como la de 2 capas.

Estas capas consisten en una capa de presentación, una de procesos y otra de datos. Lo normal es que la capa de procesos se sitúe entre la capa de datos y la de interfaz para deslocalizar los procesos del interfaz y de la capa de datos. Esto permite integrar múltiples fuentes [66].

Esta arquitectura multinivel separa efectivamente la lógica de diseño de la lógica usada para el negocio. Es por ello que si se necesitan hacer cambios en algún punto del diseño no es necesario hacer cambios en la parte de negocio.

Las capas se definen a continuación. Éstas son:

- **Capa de presentación:** esta capa es la encargada de comunicar y mostrar al usuario la información de la página. También tiene la misión de capturar toda la información que éste quiera introducir a la página.

Esta parte suele conocerse como la interfaz gráfica para el usuario. Esta interfaz debe ser intuitiva y fácil de utilizar. Como se ha mencionado anteriormente esta capa se comunica únicamente con la capa de negocio, no tendrá contacto directo con la capa de datos.

En el caso del sistema desarrollado se le presentarán los distintos enlaces de navegación entre las páginas así como la información sobre los términos que quiera consultar.

El usuario puede interactuar con el sistema a través de un cuadro de texto donde puede escribir la consulta que desea. La respuesta a esta consulta se representará en una gráfica.

- **Capa de negocio:** en esta capa se reciben los datos que se envían desde la capa de presentación. En esta capa se encuentran los métodos que se ejecutan y contienen toda la lógica de negocio, con el orden de transacciones pertinente para estas operaciones

Las respuestas de estas operaciones se vuelven a enviar a la capa de presentación, si así se requiere, o por el contrario van a la capa de datos para ser guardadas dentro de la base de datos.

Esta capa reside dentro del servidor web, por lo que está en contacto con la capa de datos a la cual hará peticiones de datos e introducirá información en la misma cuando el administrador se lo ordene.

- **Capa de datos:** en esta capa se almacenan todos los datos necesarios para la ejecución normal del sistema. Esta capa está encargada de acceder a toda esa información para facilitarla a la capa de negocio. Esta capa tiene un gestor de bases de datos que se ocupa del almacenamiento de información, solicitudes de almacenar datos y consultas de datos.

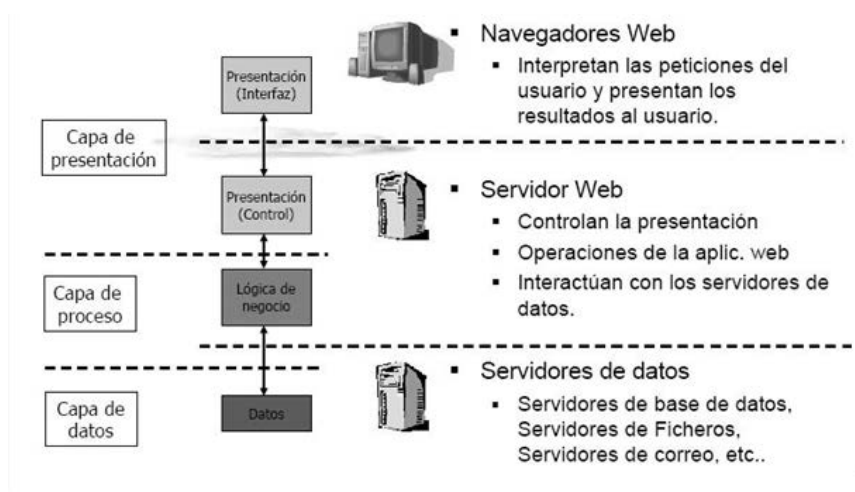


Ilustración 11. Capas de los sistemas. Fuente [66]

4.1.1.2 Estándares y normas de diseño.

En el desarrollo de la codificación del sistema se han seguido una serie de estándares para mantener en la mayor medida de lo posible la homogeneidad del código desarrollado. Para ello se han establecido las siguientes normas:

- **Clases:** se recomienda que los nombres de las clases deben empezar en mayúscula y continuar en minúscula. Se pueden alternar letras mayúsculas en el caso de que se combinen dos palabras en una.
- **Métodos:** se recomienda que el método empiece en minúscula, en el caso de que se concatenen dos o más palabras se pueden alternar letras mayúsculas. Se recomienda que estos nombres sean descriptivos.
- **Constantes:** las constantes se escriben en letras minúsculas pudiendo alternar alguna letra mayúscula si se combinan dos o más palabras. Se recomienda que sean nombres orientativos.
- **Atributos:** se escriben en letras minúsculas pudiendo alternar alguna letra mayúscula si se combinan dos o más palabras.

No es obligatorio tener que escribir comentarios en todos los métodos del sistema, no obstante se recomienda como buena práctica. Estos comentarios sirven de ayuda para clarificar la funcionalidad del sistema en el futuro y facilitar así las tareas de mantenimiento. Esta decisión de poner los comentarios está bajo la voluntad del desarrollador.

4.2 La interfaz de usuario

La interfaz de usuario que se utiliza para el sistema, es clara sencilla y con pocos elementos decorativos para asegurar que la carga de la página sea lo más ligera posible. El interfaz principal es común para todas las demás páginas, lo que cambia son los botones y cuadros de texto que se necesitan para representar la distinta información que se debe presentar al usuario.

4.3 Identificación y especificación de clases

Tras haber definido los casos de uso que van a componer el sistema, se procede a identificar los objetos y clases necesarios para desarrollar las funcionalidades especificadas en los casos de uso.

Es por ello que hay varios tipos de clases que se presentan a continuación:

- **Clases de entidad:** encargadas de representar los datos contenidos dentro de la base de datos.
- **Clases de apoyo a la interfaz:** encargadas de representar información creada por la lógica de negocio, que tiene el objetivo de ser representada.
- **Clases de interfaz:** encargadas de establecer la comunicación visual con el actor.
- **Clases de control:** encargadas de gestionar la lógica de negocio y el orden de la secuencia de ejecución de las transacciones.

Para cada tipo se escribirá un apartado en el que se identificarán las clases con sus respectivas propiedades. Se describen las funciones y los atributos que tiene cada clase para que se cumplan con los requisitos especificados. Pag 95 y 133

4.3.1.1 Clases de entidad.

Las clases entidad encargadas de representar toda la información presente en la base de datos MariaDB del sistema. Estas clases deben de tener incluida la librería java.sql.*. Estas clases son:

Nombre	GuardaDatos
Objetivos:	<p>Debe transmitir la información desde la base de datos del sistema al controlador de la lógica del negocio.</p>
Métodos y operaciones:	<pre>public void insertarDatosFacebookRaw(String idPost, String fechaPost, String messagePosts)</pre> <p>Método que inserta los comentarios (posts) que se recogen de las distintas páginas de Facebook dentro de la base de datos del sistema.</p> <pre>public void borrarDatosFacebookRaw()</pre> <p>Método que borra todos los comentarios que haya guardados en la base de datos.</p> <pre>public ArrayList<Comentario> seleccionarTodoDeDatosFacebookRaw()</pre> <p>Método que inserta todos los comentarios (posts) que hay en la base de datos y los guarda en un ArrayList de la clase Comentario.</p> <pre>public void tf(ArrayList<Comentario> lista)</pre> <p>Método que determina la frecuencia TF, el número de veces que una palabra relevante "d" se repite en el texto "t".</p> <pre>public ArrayList<ListaComidas> conexionDiccionarioPalabrasConArray()</pre> <p>Método que selecciona todas las palabras relevantes (términos clave) y los inserta en un ArrayList con una clase especial llamada ListaComidas.</p> <pre>public ArrayList<PalabrasNegativas> conexionDiccionarioPNegativasConArray()</pre> <p>Método que selecciona todas las palabras negativas y las inserta en un ArrayList con una clase especial llamada PalabrasNegativas.</p> <pre>public void insertardocumentoSoloIncognitasDB(ArrayList<Comentario> lista)</pre> <p>Método destinado a verificar con Weka, la relevancia y el porcentaje de aciertos que tiene el método tf respecto al acierto del análisis de los comentarios (posts).</p> <pre>public void insertardocumentoRelevanteDB(String idPost, String relevante, double frecuencia, double longiTexto, double palabrasPenalizadoras)</pre> <p>Método que inserta el resultado del análisis de datos con los diccionarios de las palabras relevantes y las negativas que ha obtenido de explorar cada comentario contra los dos diccionarios.</p>

	<pre>public int existeRegistro(String palabra, ArrayList<ListaComidas> listaDiccionario)</pre> <p>Método que comprueba contra la base de datos de las palabras relevantes. Esto devuelve un número entero con la frecuencia de aparición para ver si la palabra que ha introducido el cliente existe en el diccionario.</p> <pre>public ArrayList<Comentario> joinDatosFacebookRawYAnalisisrelevancia()</pre> <p>Método que selecciona los posts que han sido catalogados como relevantes por el sistema usando el algoritmo tf. Inserta los comentarios en un ArrayList.</p> <pre>public PalabraYFrecuencias frecuenciaPalabraPorMes(ArrayList<Comentario> lista, String palabraBuscada)</pre> <p>Método que en cada comentario relevante cuenta cuantas veces aparece una palabra relevante en cada mes del año.</p>
<p>Atributos</p>	<p>En esta clase existen un conjunto muy grande de atributos por lo que solo se mencionarán los más relevantes.</p> <p>Connection con; Driver para fabricar la conexión con la base de datos.</p> <p>PreparedStatement statement; Sentencia para ejecutar la consulta SQL contra la base de datos.</p> <p>ResultSet resultado; Recoge los datos que salgan de la consulta.</p> <p>ArrayList<Comentario> lista; Es un ArrayList que recoge todos los comentarios de las páginas de Facebook.</p> <p>ArrayList<ListaComidas> diccionario; Recoge en un ArrayList todas las palabras relevantes para crear así un diccionario de palabras relevantes dentro de un ArrayList.</p> <p>ArrayList<PalabrasNegativas> diccionarioNegativo; Recoge en un ArrayList todas las palabras negativas para crear así un diccionario de palabras negativas dentro de un ArrayList.</p>

Tabla 42- Guarda datos

4.3.1.2 Clases de apoyo a la interfaz

Las clases que sirven para representar gráficamente la información en la base de datos con ayuda de el módulo controlador que contiene la lógica de negocio.

Nombre	simpleline
Objetivos:	Debe graficar los resultados de las frecuencias obtenidas a lo largo de los 12 meses del año. Encargada de crear la tabla que luego se mostrará en un JSP.
Métodos y operaciones:	<p>public BufferedImage representarTablaMes(PalabraYFrecuencias frecuenciaMes)</p> <p>Método que grafica los resultados de frecuencia de la palabra seleccionada por cada mes.</p>
Atributos:	<p>BufferedImage img; Es un objeto para representar la imagen renderizada por el método.</p> <p>PalabraYFrecuencias frecuenciaMes; Es un objeto especial para poder contener toda la información que se necesita para representar la tabla, como las frecuencias de cada mes y la palabra en concreto.</p>

Tabla 43- Simpleline

4.3.1.3 Clases de interfaz.

Estas clases presentan la información al usuario y permiten también que el usuario interactúe con el sistema.

Nombre	index
Objetivos:	Genera un índice inicial para que el usuario pueda seleccionar entre varias opciones de búsqueda sobre el campo que desee encontrar. Estas opciones se dividen entre seleccionar alimentos, bebidas, regiones, términos especiales y comidas diarias
Atributos	<p>Los métodos que posee son llamadas al método doGet de la clase controladora para redirigir la petición del usuario a la página correspondiente mediante la etiqueta:</p> <pre>Ir a la página de</pre> <pre>
</pre> <p>Con cada una de las secciones. También tiene otra para recargar el índice.</p>

Tabla 44- Index

Nombre	paginaAlimentación
Objetivos:	Esta página muestra todos los alimentos que se pueden consultar en la base de datos. Están presentados en una tabla con una caja de texto donde el usuario puede introducir el término que quiera buscar.
Atributos	<p>El usuario introduce un determinado texto en la caja de texto reservada para ello, esta consulta se lanzará contra la base de datos. La sentencia a continuación crea esa caja de texto.</p> <pre><input type="text" maxlength="20" size="20" name="busqueda"/> <input type="submit" value="Buscar"/></pre> <p>El texto que el usuario introduzca entrará por el método "doPost" de la clase controladora.</p>

Tabla 45- PáginaAlimentación

Nombre	Ejemplo6 (Bebidas)
Objetivos:	Esta página muestra todos las bebidas que se pueden consultar en la base de datos. Están presentados en una tabla con una caja de texto donde el usuario puede introducir el término que quiera buscar.
Atributos	<p>El usuario introduce un determinado texto en la caja de texto reservada para ello, esta consulta se lanzará contra la base de datos. La sentencia a continuación crea esa caja de texto.</p> <pre><input type="text" maxlength="20" size="20" name="busqueda"/> <input type="submit" value="Buscar"/></pre> <p>El texto que el usuario introduzca entrará por el método "doPost" de la clase controladora.</p>

Tabla 46- Ejemplo6(bebidas)

Nombre	paginaRegiones
Objetivos:	Esta página muestra todos las regiones que se pueden consultar en la base de datos. Están presentados en una tabla con una caja de texto donde el usuario puede introducir el término que quiera buscar.

Atributos	<p>El usuario introduce un determinado texto en la caja de texto reservada para ello, esta consulta se lanzará contra la base de datos. La sentencia a continuación crea esa caja de texto.</p> <pre><input type="text" maxlength="20" size="20" name="busqueda"/> <input type="submit" value="Buscar"/></pre> <p>El texto que el usuario introduzca entrará por el método “doPost” de la clase controladora.</p>
------------------	---

Tabla 47-PáginaRegiones

Nombre	paginaTerminosEspeciales
Objetivos:	Esta página muestra todos los términos especiales que se pueden consultar en la base de datos. Están presentados en una tabla con una caja de texto donde el usuario puede introducir el término que quiera buscar.
Atributos	<p>El usuario introduce un determinado texto en la caja de texto reservada para ello, esta consulta se lanzará contra la base de datos. La sentencia a continuación crea esa caja de texto.</p> <pre><input type="text" maxlength="20" size="20" name="busqueda"/> <input type="submit" value="Buscar"/></pre> <p>El texto que el usuario introduzca entrará por el método “doPost” de la clase controladora.</p>

Tabla 48- PáginaTérminosEspeciales

Nombre	paginaTiposComida
Objetivos:	Esta página muestra todas las comidas diarias que se pueden consultar en la base de datos. Están presentados en una tabla con una caja de texto donde el usuario puede introducir el término que quiera buscar.
Atributos	<p>El usuario introduce un determinado texto en la caja de texto reservada para ello, esta consulta se lanzará contra la base de datos. La sentencia a continuación crea esa caja de texto.</p> <pre><input type="text" maxlength="20" size="20" name="busqueda"/> <input type="submit" value="Buscar"/></pre> <p>El texto que el usuario introduzca entrará por el método “doPost” de la clase controladora.</p>

Tabla 49-PáginaTiposComida

Nombre	presentacionTabla
Objetivos:	Esta página muestra la gráfica ya renderizada con los datos que se han extraído de la base de datos gracias al análisis de los comentarios.
Métodos y operaciones	Esta página muestra una imagen de una gráfica renderizada que ha construido la clase simpleline. Esta imagen se muestra al usuario por la pantalla para que pueda consultar la información.
Atributos	Object imgObj; Es un objeto para capturar la imagen renderizada que se va a presentar después un objeto de tipo Image. Image tabla; Es un objeto de tipo imagen que se usa para representar la tabla de forma gráfica en el interfaz.

Tabla 50-PresentaciónTabla

4.3.1.4 Clases de control.

Las clases de control son las encargadas de ejecutar la lógica de negocio y ejecutar correctamente la secuencia de las transacciones.

Nombre	ServletFormulario
Objetivos:	Esta clase controla el flujo de información entre el interfaz del usuario, la base de datos y las distintas transacciones que se producen dentro del sistema.
Métodos y operaciones:	<p>public void init() throws ServletException</p> <p>Este método inicializa varios objetos, así como dos ArrayList de palabras negativas y positivas para que puedan ser usados por el sistema durante su ejecución.</p> <p>protected void doGet(HttpServletRequest request, HttpServletResponse response)</p> <p>Método que se encarga de responder a las peticiones que vienen por doGet, estas peticiones de usuario suelen ser peticiones para acceder a otras páginas. Estas peticiones a su vez son gestionadas por el método processRequest.</p> <p>protected void processRequest(HttpServletRequest request, HttpServletResponse response)</p>

	<p>Este método se encarga de procesar cada petición que llega desde el método doGet para redirigir a donde sea necesario al usuario en la navegación por las páginas.</p> <pre>protected void doPost(HttpServletRequest request, HttpServletResponse response) throws ServletException, IOException</pre> <p>Este método gestiona de responder a las peticiones de usuario que vienen por doPost. Estas peticiones suelen traer contenido de tipo texto plano por lo que deben ser analizadas.</p> <pre>protected void procesoPost(HttpServletRequest request, HttpServletResponse response) throws ServletException, IOException</pre> <p>Este método procesa toda la información que le llega desde doPost para llamar a los métodos que analizan los comentarios (posts) y crean las gráficas que luego se muestran al usuario.</p>
Atributos	<p>ArrayList<Comentario> lista; crea una lista de todos los posts que hay guardados en la base de datos.</p> <p>ArrayList<ListaComidas> listaDiccionario; crea una lista con todas las palabras relevantes que se tienen en el diccionario de palabras relevantes dentro de la base de datos.</p>

Tabla 51- ServletFormulario

Nombre	Comentario
Objetivos:	Esta clase es la encargada de recoger los comentarios (posts) de la red social identificando 3 atributos de interés.
Métodos y operaciones:	<pre>public String getIdPost(); Retorna el identificador del comentario (post)</pre> <pre>public String getFechaPost(); Retorna la fecha del comentario (post)</pre> <pre>public String getTexto(); Retorna la fecha del comentario (post)</pre> <pre>public void setIdPost(String id); Modifica el identificador del comentario (post) con el parámetro que recibe.</pre> <pre>public void setFechaPost(String date); Modifica la fecha del comentario (post) con el parámetro que recibe.</pre> <pre>public void setTextoPost(String txt); Modifica el texto del comentario (post) con el parámetro que recibe.</pre>

Atributos:	String idPost; Recoge el Identificativo para el comentario (post) String fecha; Recoge la fecha del comentario (post) String texto; Recoge el texto del comentario.
-------------------	---

Tabla 52- Comentario

Nombre	ListaComidas
Objetivos:	Esta clase representa al diccionario de sinónimos con sus cuatro subcategorías para catalogar claramente a cada término.
Métodos y operaciones:	<p>public String getCat1(); Retorna el valor de la primera subcategoría. Este método es exactamente igual para las otras 3 subcategorías.</p> <p>public void setCat1(String cat); Modifica el valor de la primera categoría. Este método es exactamente igual para las otras 3 subcategorías.</p>
Atributos:	<p>Cada uno de los string representa un nivel de categoría y los subniveles de este:</p> <p>String categoria1; Representa la categoría general String categoria2; Representa el tipo de alimento String categoria3; Representa el género String categoria4; Representa los sub géneros o marcas.</p>

Tabla 53- ListaComidas

Nombre	PalabraYFrecuencias
Objetivos:	Esta clase registra las frecuencias de uso de una palabra concreta en cada mes del año. Así se puede representar después en una gráfica.
Métodos y operaciones:	<p>public String getPalabraClave() Retorna el valor de la palabra clave para la cual se está ejecutando el análisis.</p> <p>public void setPalabraClave(String palabraClave); Modifica el valor de la palabra clave para la cual se ha ejecutado el análisis.</p> <p>Para los valores de frecuencia de los meses del año se han creado 12 métodos muy parecidos entre sí, que solo difieren en el mes para el que va destinado la frecuencia. Por lo que solo se explicará el primero.</p> <p>public int getEnero(); Retorna el valor de la frecuencia de la palabra buscada en el mes de enero.</p> <p>public void setEnero(int enero); Modifica el valor de la frecuencia para el mes de enero para la palabra buscada.</p>

Atributos:	<p>String palabraClave; Es el término clave del cual se está analizando su frecuencia de uso.</p> <p>Se usa un registro de tipo int para registrar las frecuencias en cada mes del año.</p> <p>int enero; int febrero; int marzo; int abril; int mayo; int junio; int julio; int agosto; int septiembre; int octubre; int noviembre; int diciembre;</p>
-------------------	---

Tabla 54- PalabrayFrecuencias

Nombre	PalabrasNegativas
Objetivos:	Esta clase almacena una palabra negativa que corresponde al diccionario de palabras negativas que se usa para el análisis.
Métodos y operaciones:	<p>public String getPalabraNegativa(); Recupera el valor de la palabra negativa que se ha guardado.</p> <p>public void setPalabraNegativa(String palabN); Modifica el valor de la palabra negativa del diccionario de palabras negativas.</p>
Atributos:	String palabraNeg; guarda el valor de la palabra negativa para el diccionario de palabras negativas.

Tabla 55- PalabrasNegativas

Nombre	TextDirectoryToArff
Objetivos:	Esta clase es la encargada de generar archivos de tipo ARFF para que el sistema de análisis de minería de datos Weka pueda analizar mediante el algoritmo "PART" la eficiencia del análisis.
Métodos y operaciones:	public void conversorTextosArff(); Este método convierte los archivos de texto de la base de datos en archivos del tipo ARFF para que puedan ser analizados por Weka. Se puede usar cualquier carpeta del equipo a la que no haya restricciones de uso.
Atributos:	<p>Instances data; guarda los archivos de texto recuperados del directorio del equipo.</p> <p>ArffSaver saber; encargada de almacenar los archivos que se convierten al tipo ARFF.</p>

Tabla 56- TextDirectoryToArff

Nombre	WriteFile
Objetivos:	Esta clase es la encargada de escribir los textos que se han almacenado en la base de datos dentro de un directorio en el equipo para que estos luego sean convertidos y analizados en Weka.
Métodos y operaciones:	public void crearFicheroDeTexto(ArrayList<Comentario> lista); Este método guarda los comentarios en ficheros de texto dentro de un directorio especificado.
Atributos:	int tam; guarda el tamaño de la lista de comentarios (posts) que vienen por parámetro Writer writer; Escribe el texto del comentario dentro de un archivo de texto en el directorio especificado. String id; Guarda el id del comentario para añadirlo como nombre del archivo de texto. String fecha; Guarda la fecha en la que se realizó el comentario para después añadirla dentro del archivo de texto. String texto; Guarda el texto del comentario para luego guardarlo dentro del archivo de texto.

Tabla 57- WriteFile

Diagrama de clases. La información de las distintas clases del diagrama se encuentra en las tablas especificadas en las páginas anteriores por razones de claridad, ya que de otra forma no se verían de forma nítida en éste diagrama.

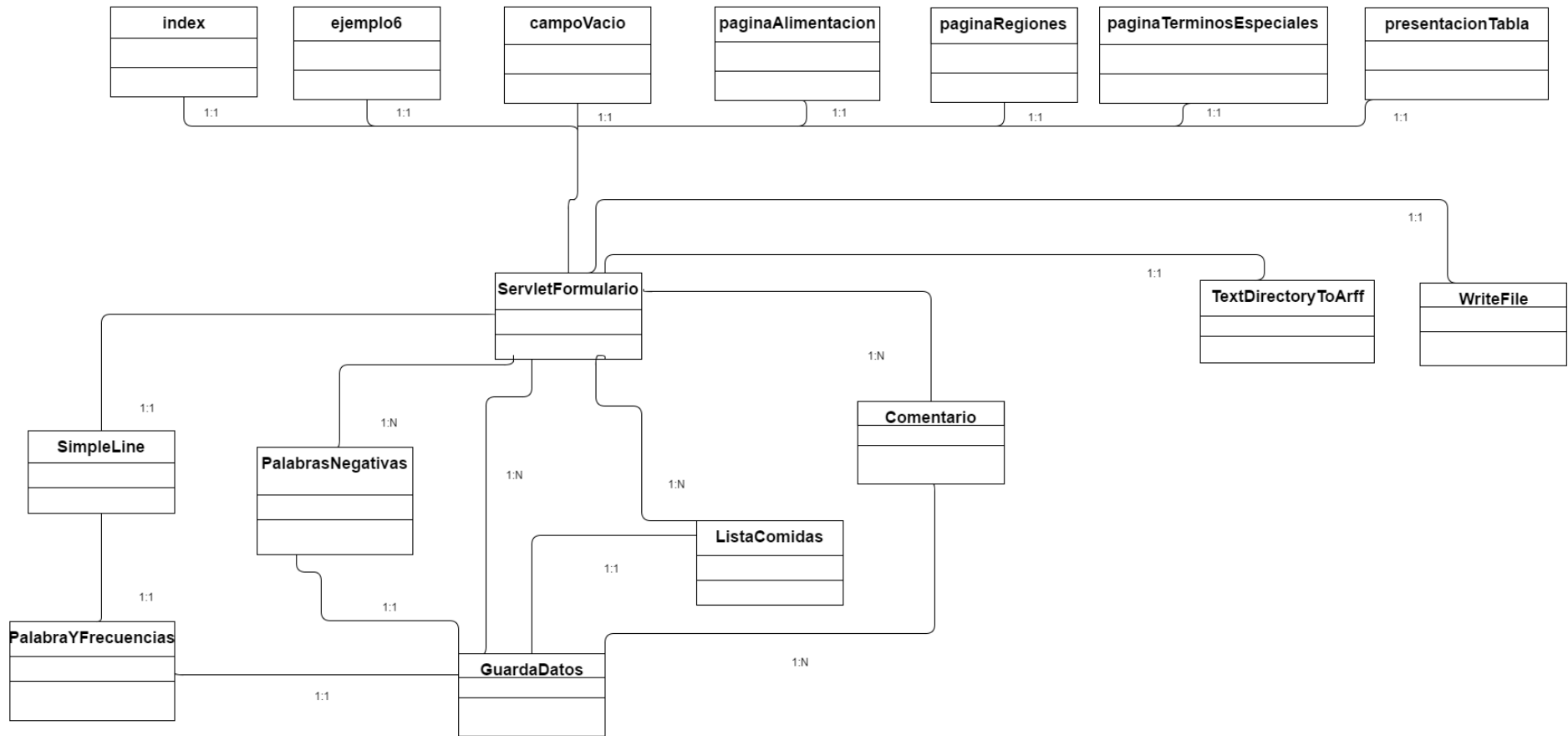


Ilustración 12. Diagrama de clases. Fuente: elaboración propia

4.4 Diseño físico de datos

Se procede a explicar el diseño físico de la base de datos. Este estudio se hace con la intención de clarificar la estructura a la vez que se garantiza una integridad de datos a través de evitar las redundancias innecesarias. 158 107

Se describirán las entidades que forman la base de datos, este modelo de datos se ha elaborado antes de construir la base de datos, esto sirve para poder organizar todos los elementos de información que se encuentran en dicha base de datos de la forma más eficiente.

Las entidades a explicar son:

- **Postsfacebookraw:** esta tabla almacena todos los comentarios (posts) que se descargan de la red social.
- **Diccionariopalabras:** esta tabla almacena las palabras relevantes o términos clave.
- **Palabrasnegativas:** esta tabla almacena las palabras negativas para el análisis.
- **Postrelevancia:** esta tabla almacena los resultados de los comentarios analizados guardando el id del comentario y los datos relativos a su relevancia.
- **Postrelevanciaincognitas:** esta tabla almacena la misma información que la tabla postrelevancia con la salvedad de que el campo que indica si es relevante o no ha sido rellenado con una incógnita. Esta se utiliza exclusivamente para hacer las pruebas con Weka.

Cabe mencionar que el campo idPost es la clave primaria, esta se puede escribir como primary key o PK. La cual identifica de forma única a cada fila de una tabla. Por lo que no puede haber dos filas en una tabla que tengan la misma clave primaria [67].

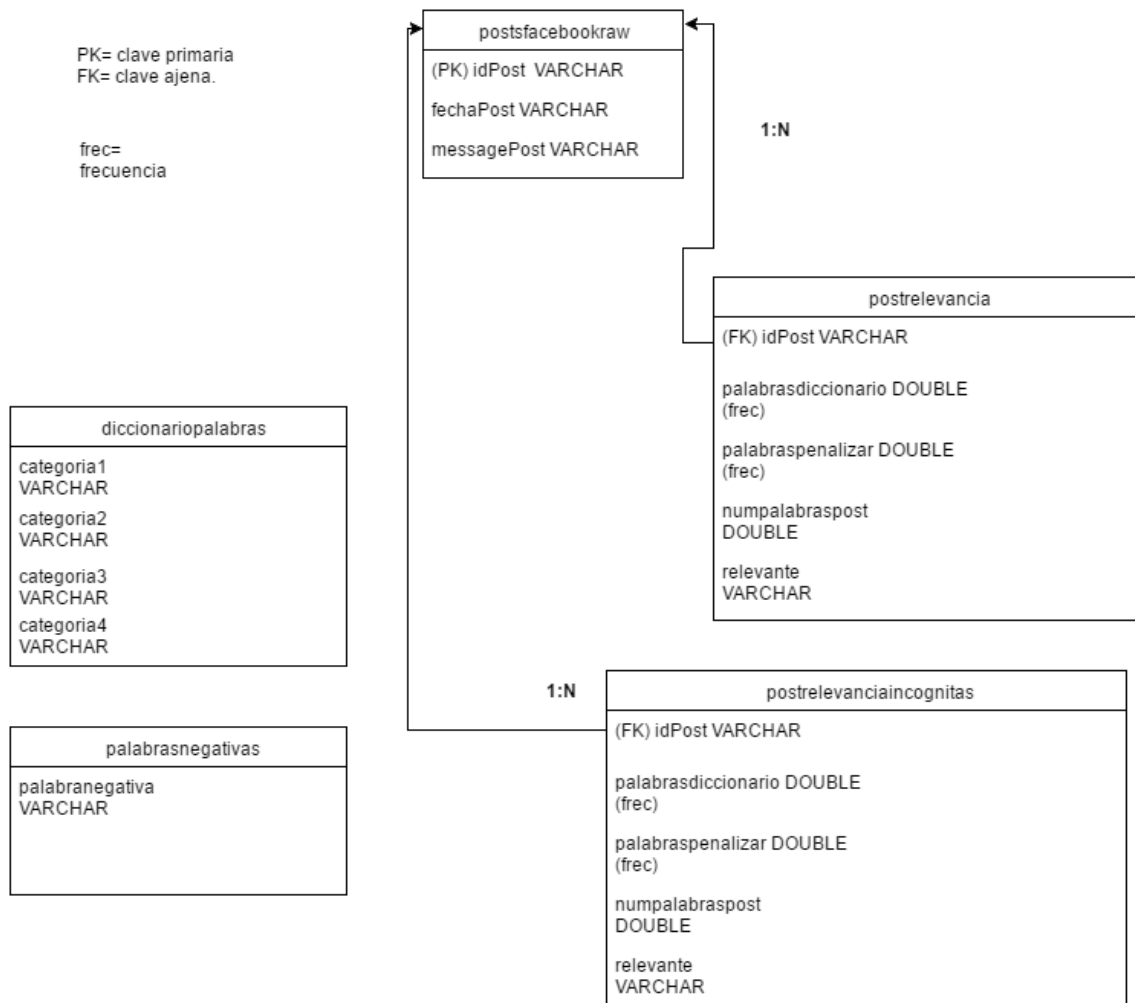


Ilustración 13. Diseño de datos. Fuente: elaboración propia

4.5 Migración de datos inicial.

Para poder llevar a cabo las tareas básicas de búsqueda de frecuencias y cálculo del algoritmo TF es necesario llevar a cabo una carga inicial de datos. Esta carga no se puede hacer de cualquier forma ya que debe cumplir con unos requisitos previos para poder hacerla.

Al realizar borrados, o modificaciones en las tablas de datos no se debe alterar la estructura de estas tablas. En cualquier caso junto a la entrega de la aplicación se facilitará una copia de un estado valido para la base de datos.

4.5.1 El entorno

Para poder hacer un uso del sistema en condiciones ideales es necesario tener instalado en el equipo el sistema XAMPP. Este sistema provee el un gestor de bases de datos que se llama MariaDB. Se recomienda usar una versión 5.6 o posterior de XAMPP.

4.5.2 Procedimiento de carga inicial de datos

Para que el sistema funcione correctamente, se requiere tener cargados los datos en la base de datos. Esto es necesario ya que de lo contrario se tendrían que modificar las tablas de la base de datos manualmente cada vez que se desplegase la aplicación.

Se explica a continuación el proceso de carga para los datos en el sistema de XAMPP, que en este caso pes MariaDB.

4.5.2.1.1 1º Acceder a phpMyAdmin

Debemos acceder a la página inicial del sistema gestor, en este caso es phpMyAdmin. Es necesario activar en el panel de control de XAMPP los servicios de MYSQL y Apache.

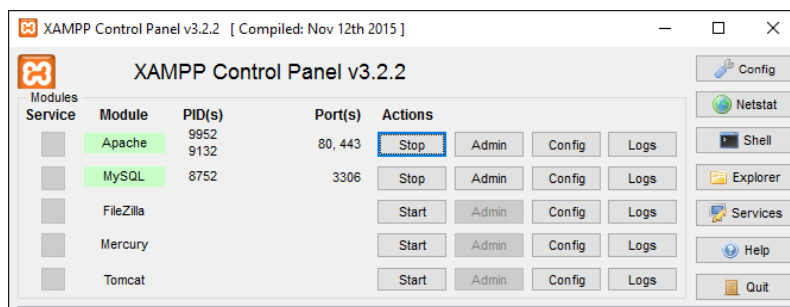


Ilustración 14. Panel de control XAMPP. Fuente: elaboración propia

Acceder desde el navegador a la URL del phpMyAdmin para ello debemos escribir <http://localhost/phpmyadmin/> en el navegador:

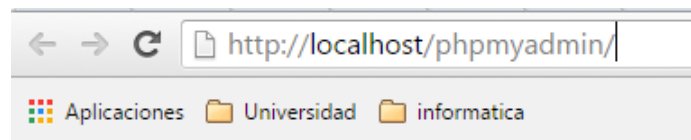


Ilustración 15. Demostración de URL. Fuente: elaboración propia

4.5.2.1.2 2º Cargar la copia de seguridad de la base de datos.

Se debe cargar el archivo .sql para cargar toda la información necesaria para hacer funcionar la aplicación. Para ello seleccionamos la opción Importar dentro del menú principal de phpMyAdmin.

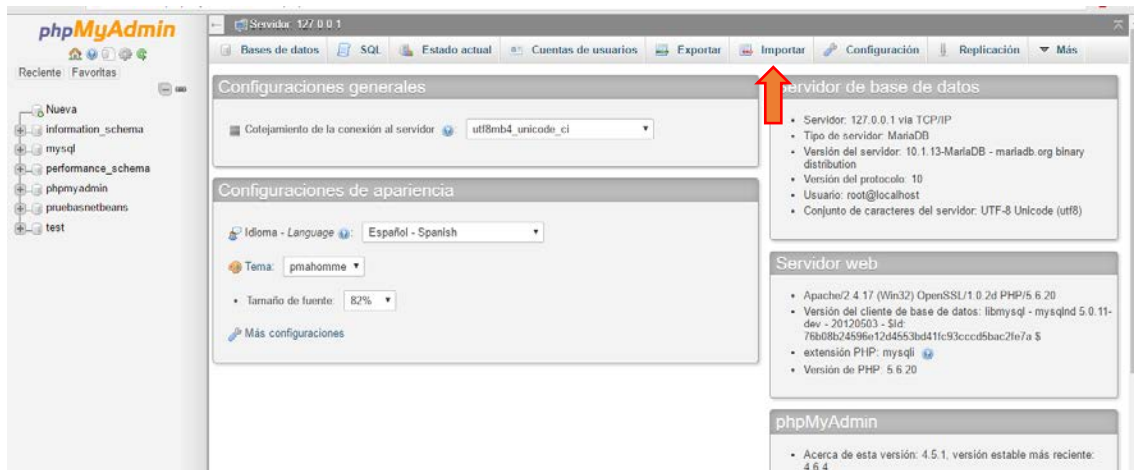


Ilustración 16. PHPmyAdmin. Fuente: elaboración propia

Acto seguido se pincha en la opción de seleccionar archivo. Es importante marcar que el conjunto de caracteres del archivo sea iso-8859-1 como se puede observar en la imagen.

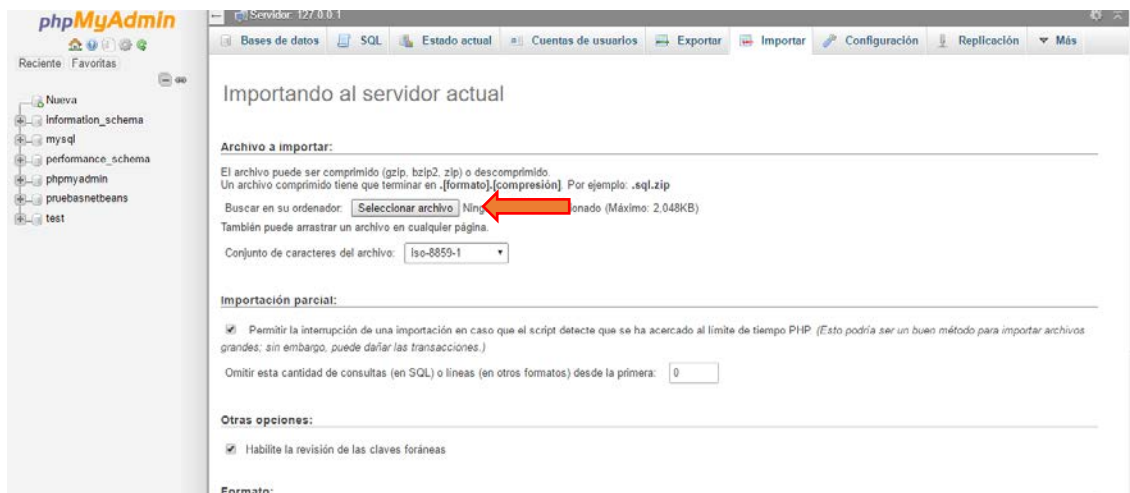


Ilustración 17. Subir datos a XAMPP. Fuente: elaboración propia

4.5.2.1.3 3º Elegir ruta con la copia de la base de datos.

Se debe seleccionar la ruta donde se encuentra la copia de la base de datos con el archivo .sql en el equipo.

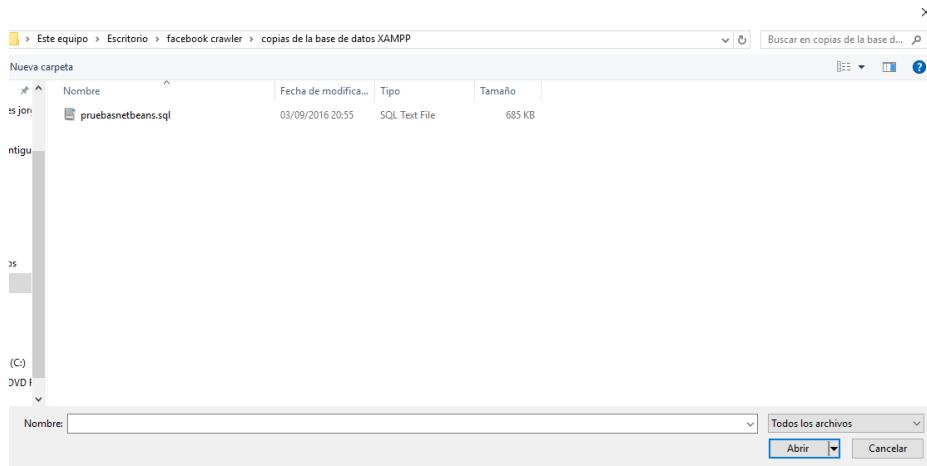


Ilustración 18. Ver fichero de carga. Fuente: elaboración propia

Después de haberlo seleccionado, en la pantalla del navegador deslizaremos hacia abajo la página hasta encontrar el botón de continuar. Una vez presionado el botón la base de datos se cargará, esto puede llevar unos minutos, se ruega ser paciente.

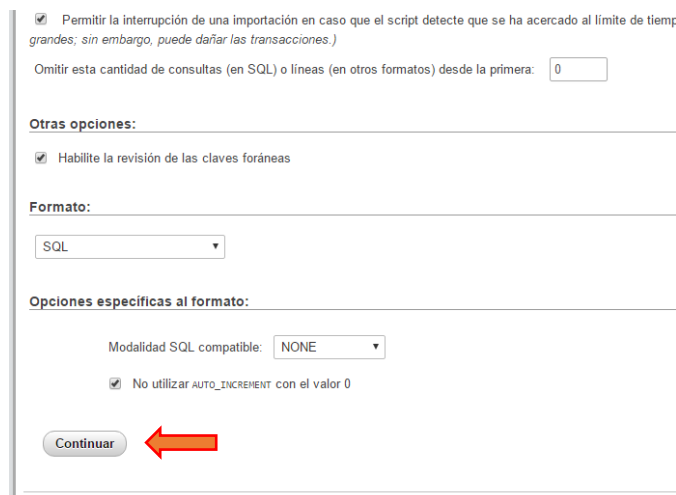


Ilustración 19. Selección de carga de tablas. Fuente: elaboración propia

5 Especificación técnica del plan de pruebas

El plan de pruebas consiste en un documento en el cuál se especifica cómo se llevarán a cabo las pruebas del sistema con la meta de proporcionar información objetiva sobre la calidad del sistema a los stakeholders. Estos stakeholders son el cliente que en este caso está representado por el tribunal de corrección de la UC3M.

Por la naturaleza y el tipo de prueba ésta podrá ser realizada en un determinado momento del proyecto. El objetivo es que el plan tenga más probabilidades de encontrar los defectos con el menor esfuerzo posible.

A continuación se explica la tabla que se va a utilizar para describir las pruebas:

Nombre	P-ZN
Elementos de prueba	
Descripción:	
Especificaciones de entrada:	
Especificaciones de salida:	
Requisitos especiales del procedimiento:	

Tabla 58- PZN

El formato de la tabla de prueba contiene la siguiente información:

- Donde el nombre de la prueba se compone de:
 - P que indica que es un caso de prueba.
 - Z que indica el tipo de prueba que es.
 - N que es el número de prueba que es.
- Elemento de la prueba: indica el requisito probado si es que lo hay.
- Descripción: descripción de la prueba.
- Especificaciones de entrada: datos necesarios que se deben ingresar en el sistema para realizar la prueba.
- Especificaciones de salida: respuesta del sistema esperada.
- Requisitos especiales de procedimiento: condiciones previas que se necesitan para hacer la prueba.

5.1 Pruebas funcionales

La herramienta debe cumplir todas las funciones especificadas en los requisitos de usuario.

Nombre	P-F01
Elementos de prueba	RF-01
Descripción:	Se deben cargar los datos de los comentarios (posts) de Facebook en la base de datos.
Especificaciones de entrada:	Se deben cargar los datos de los comentarios desde la propia API de Facebook y con ayuda del código implementado. También se pueden cargar desde el archivo .sql que se haya entregado.
Especificaciones de salida:	Se apreciará en el log de la base de datos que se han cargado los datos satisfactoriamente.

Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y opcionalmente el fichero .sql con una copia de la base de datos.
Resultado obtenido:	Satisfactorio

Tabla 59– PF01

Nombre	P-F02
Elementos de prueba	RF-02
Descripción:	Se deben cargar los datos de las palabras del diccionario de palabras relevantes en el campo de la alimentación.
Especificaciones de entrada:	Se debe disponer del fichero .csv o preferiblemente el archivo .sql para tener cargados todos los términos.
Especificaciones de salida:	Se apreciará en el log de la base de datos que se han cargado los datos satisfactoriamente.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y opcionalmente el fichero .sql con una copia de la base de datos un fichero .csv.
Resultado obtenido:	Satisfactorio

Tabla 60– PF02

Nombre	P-F03
Elementos de prueba	RF-03
Descripción:	Se deben cargar los datos de las palabras del diccionario de palabras negativas.
Especificaciones de entrada:	Se debe disponer del fichero .csv o preferiblemente el archivo .sql para tener cargados todos los términos.
Especificaciones de salida:	Se apreciará en el log de la base de datos que se han cargado los datos satisfactoriamente.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y opcionalmente el fichero .sql con una copia de la base de datos un fichero .csv.
Resultado obtenido:	Satisfactorio

Tabla 61–PF03

Nombre	P-F04
Elementos de prueba	RF-04
Descripción:	El sistema debe ser capaz de mostrar en una gráfica las frecuencias de uso de la palabra que el usuario ha introducido y desea conocer los resultados.
Especificaciones de entrada:	El usuario debe haber entrado en la página y haber escrito en el cuadro de texto el término de búsqueda que se encuentra en el listado ofrecido.
Especificaciones de salida:	Se mostrará una gráfica con el listado de frecuencias.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 62-PF04

Nombre	P-F05
Elementos de prueba	RF-05
Descripción:	El usuario debe poder introducir uno de los términos deseados en el cuadro de texto de la página.
Especificaciones de entrada:	El usuario debe haber entrado en la página dentro de la categoría que haya seleccionado.
Especificaciones de salida:	A continuación de enviar la petición formulada con la palabra deseada se mostrará otra página.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 63-PF05

Nombre	P-F06
Elementos de prueba	RF-06
Descripción:	Se deben mostrar al usuario las palabras relevantes de la categoría alimentación en una tabla y con un cuadro de texto para que este pueda escribir su consulta.
Especificaciones de entrada:	El usuario debe haber seleccionado en la página de inicio el enlace de la categoría de Alimentación.
Especificaciones de salida:	Debe cargarse toda la página.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 64-PF06

Nombre	P-F07
Elementos de prueba	RF-07
Descripción:	Se deben mostrar al usuario las palabras relevantes de la categoría bebidas en una tabla y con un cuadro de texto para que este pueda escribir su consulta.
Especificaciones de entrada:	El usuario debe haber seleccionado en la página de inicio el enlace de la categoría de Bebidas.
Especificaciones de salida:	Debe cargarse toda la página.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 65–PF07

Nombre	P-F08
Elementos de prueba	RF-08
Descripción:	Se deben mostrar al usuario las palabras relevantes de la categoría Regiones en una tabla y con un cuadro de texto para que este pueda escribir su consulta.
Especificaciones de entrada:	El usuario debe haber seleccionado en la página de inicio el enlace de la categoría de Regiones.
Especificaciones de salida:	Debe cargarse toda la página.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 66–PF08

Nombre	P-F09
Elementos de prueba	RF-08
Descripción:	Se deben mostrar al usuario las palabras relevantes de la categoría Términos especiales en una tabla y con un cuadro de texto para que este pueda escribir su consulta.
Especificaciones de entrada:	El usuario debe haber seleccionado en la página de inicio el enlace de la categoría de Términos especiales.
Especificaciones de salida:	Debe cargarse toda la página.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 67–PF09

Nombre	P-F10
Elementos de prueba	RF-10
Descripción:	Se deben mostrar al usuario las palabras relevantes de la categoría Comidas diarias en una tabla y con un cuadro de texto para que este pueda escribir su consulta.
Especificaciones de entrada:	El usuario debe haber seleccionado en la página de inicio el enlace de la categoría de Comidas diarias.
Especificaciones de salida:	Debe cargarse toda la página.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 68–PF10

Nombre	P-F11
Elementos de prueba	RF-11
Descripción:	El sistema debe determinar la relevancia o no de un comentario (post) para su posterior análisis.
Especificaciones de entrada:	El usuario debe haber introducido un término válido en el cuadro de texto de consulta. Se deben de haber cargado los comentarios (posts) en la base de datos.
Especificaciones de salida:	El sistema debe ejecutar correctamente el análisis del comentario (post) y guardarlo en la base de datos.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

Tabla 69–PF11

Nombre	P-F12
Elementos de prueba	RF-12
Descripción:	El sistema debe analizar los comentarios (posts) que hayan sido identificados como relevantes para su análisis.
Especificaciones de entrada:	El sistema debe recibir un término válido de análisis. Deben de analizarse solo los comentarios que hayan sido catalogados como relevantes.
Especificaciones de salida:	El sistema debe ejecutar correctamente el análisis del comentario (post) y mostrarlo al usuario por pantalla en una gráfica.
Requisitos especiales del procedimiento:	Disponer de conexión a internet, el sistema XAMPP y la base de datos cargada.
Resultado obtenido:	Satisfactorio

5.2 Pruebas de aceptación.

Las pruebas de aceptación son las que se realizan al sistema que el cliente usará. Estas pruebas van encaminadas a demostrar si se cumplen con todos los requisitos y el manual de usuario que se ha entregado. Estas pruebas se realizan una vez que se ha finalizado el sistema [68].

Nombre	P-A01
Elementos de prueba	RF-06, RF-02, RF-01, C-02
Descripción:	Se mostrarán los contenidos de la categoría general de Alimentos.
Especificaciones de entrada:	Selección del enlace que lleva a la página de Alimentos. El usuario debe haber accedido a la página de inicio.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber pinchado en el enlace de la categoría general de Alimentos.
Resultado obtenido:	Satisfactorio

Tabla 71–PA01

Nombre	P-A02
Elementos de prueba	RF-03, RF-02, RF-01, RF-13, C-01, C-05
Descripción:	El usuario debe poder acceder al índice de la página a través de su navegador, visualizando las opciones de ésta.
Especificaciones de entrada:	El usuario debe introducir la dirección correcta en al URL del navegador.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe tener conexión a internet y el servicio debe estar en línea.
Resultado obtenido:	Satisfactorio

Tabla 72–PA02

Nombre	P-A03
Elementos de prueba	RF-03, RF-02, RF-01, RF-07, C-02
Descripción:	Se mostrarán los contenidos de la categoría general de Bebidas.

Especificaciones de entrada:	Selección del enlace que lleva a la página de Bebidas. El usuario debe haber accedido a la página de inicio.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber pinchado en el enlace de la categoría general de Bebidas.
Resultado obtenido:	Satisfactorio

Tabla 73 –PA03

Nombre	P-A04
Elementos de prueba	RF-03, RF-02, RF-01, RF-08, C-02
Descripción:	Se mostrarán los contenidos de la categoría general de Regiones
Especificaciones de entrada:	Selección del enlace que lleva a la página de Regiones. El usuario debe haber accedido a la página de inicio.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber pinchado en el enlace de la categoría general de Regiones.
Resultado obtenido:	Satisfactorio

Tabla 74 –PA04

Nombre	P-A05
Elementos de prueba	RF-03, RF-02, RF-01, RF-09, C-02
Descripción:	Se mostrarán los contenidos de la categoría general de Términos especiales.
Especificaciones de entrada:	Selección del enlace que lleva a la página de Términos especiales. El usuario debe haber accedido a la página de inicio.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber pinchado en el enlace de la categoría general de Términos especiales.
Resultado obtenido:	Satisfactorio

Tabla 75 –PA05

Nombre	P-A06
Elementos de prueba	RF-03, RF-02, RF-01, RF-10, C-02
Descripción:	Se mostrarán los contenidos de la categoría general de Comidas diarias
Especificaciones de entrada:	Selección del enlace que lleva a la página de Comidas diarias. El usuario debe haber accedido a la página de inicio.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber pinchado en el enlace de la categoría general de Comidas diarias.
Resultado obtenido:	Satisfactorio

Tabla 76-PA06

Nombre	P-A07
Elementos de prueba	RF-03, RF-02, RF-01, RF-06, C-02, C-03
Descripción:	El usuario debe poder explorar la tabla de términos clave de la categoría general de Alimentos.
Especificaciones de entrada:	Selección del enlace que lleva a la página de Alimentación. El usuario debe haber accedido a la página de Alimentos.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber accedido la página de la categoría general de Alimentos.
Resultado obtenido:	Satisfactorio

Tabla 77-PA07

Nombre	P-A08
Elementos de prueba	RF-03, RF-02, RF-01, RF-07, C-02, C-03
Descripción:	El usuario debe poder explorar la tabla de términos clave de la categoría general de Bebidas.
Especificaciones de entrada:	Selección del enlace que lleva a la página de Bebidas. El usuario debe haber accedido a la página de Bebidas.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber accedido la página de la categoría general de Bebidas.
Resultado obtenido:	Satisfactorio

Tabla 78-PA08

Nombre	P-A09
Elementos de prueba	RF-03, RF-02, RF-01, RF-08, C-02, C-03
Descripción:	El usuario debe poder explorar la tabla de términos clave de la categoría general de Regiones.
Especificaciones de entrada:	Selección del enlace que lleva a la página de Regiones. El usuario debe haber accedido a la página de Regiones.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber accedido la página de la categoría general de Regiones.
Resultado obtenido:	Satisfactorio

Tabla 79–PA09

Nombre	P-A10
Elementos de prueba	RF-03, RF-02, RF-01, RF-09, C-02, C-03
Descripción:	El usuario debe poder explorar la tabla de términos clave de la categoría general de Términos especiales.
Especificaciones de entrada:	Selección del enlace que lleva a la página de Términos especiales. El usuario debe haber accedido a la página de Términos especiales.
Especificaciones de salida:	Se debe mostrar la página requerida.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a la página del índice y haber accedido la página de la categoría general de Términos especiales.
Resultado obtenido:	Satisfactorio

Tabla 80–PA10

Nombre	P-A11
Elementos de prueba	RF-03, RF-02, RF-01, RF-05, C-02, C-04
Descripción:	El usuario debe poder introducir en un cuadro de texto el término que desea buscar.
Especificaciones de entrada:	El usuario debe introducir el término que desea buscar de cualquier categoría genérica.
Especificaciones de salida:	Se debe mostrar en otra página una gráfica con los resultados del análisis sobre la palabra consultada.
Requisitos especiales del procedimiento:	El usuario debe haber accedido a una de las categorías genéricas.
Resultado obtenido:	Satisfactorio

Tabla 81-PA11

Nombre	P-A12
Elementos de prueba	RF-03, RF-02, RF-01, RF-12, C-02, C-03
Descripción:	El usuario debe poder ver la gráfica de resultados del término que ha introducido en el cuadro de texto.
Especificaciones de entrada:	El usuario debe introducir el término que desea buscar de cualquier categoría genérica. El modulo controlador captura esta palabra y procesa la petición.
Especificaciones de salida:	Se debe mostrar una página con la gráfica del término deseado.
Requisitos especiales del procedimiento:	El usuario debe haber escrito un término válido que se encuentre en alguna de las clases genéricas.
Resultado obtenido:	Satisfactorio

Tabla 82-PA12

Nombre	P-A13
Elementos de prueba	RF-03, RF-02, RF-11, RF-12, C-02, C-04
Descripción:	El usuario debe poder ver la gráfica de resultados del término que ha introducido en el cuadro de texto.
Especificaciones de entrada:	El usuario debe introducir el término que desea buscar de cualquier categoría genérica. El modulo controlador captura esta palabra y procesa la petición.
Especificaciones de salida:	Se debe mostrar una página con la gráfica del término deseado.
Requisitos especiales del procedimiento:	El usuario debe haber escrito un término válido que se encuentre en alguna de las clases genéricas.
Resultado obtenido:	Satisfactorio

Tabla 83-PA13

Nombre	P-A14
Elementos de prueba	RF-13, C-05
Descripción:	El usuario debe poder volver al inicio desde cualquier página de las categorías genéricas.
Especificaciones de entrada:	El usuario debe pinchar sobre uno de los enlaces que están colocados en la página.
Especificaciones de salida:	Se debe volver al índice.
Requisitos especiales del procedimiento:	
Resultado obtenido:	Satisfactorio

Tabla 84-PA14

5.3 Pruebas no funcionales

Nombre	P-NF01
Elementos de prueba	RNF-PE-01
Descripción:	El tiempo de respuesta debe ser menor a 10 minutos ante una consulta.
Especificaciones de entrada:	Entra una petición al módulo de control
Especificaciones de salida:	Se devuelve una respuesta acorde a la entrada y a la salida que espera el usuario.
Requisitos especiales del procedimiento:	
Resultado obtenido:	Satisfactorio

Tabla 85-PNF01

Nombre	P-NF02
Elementos de prueba	RNF-U-01, RNF-I-01, RNF-I-09
Descripción:	La interfaz debe ser clara y fácil de entender siguiendo un patrón común. Moverse de una página a otra se puede hacer con botones o enlaces.
Especificaciones de entrada:	Una orden para navegar a otra página como pinchar e un botón o en un enlace.
Especificaciones de salida:	Navegar por las páginas que el cliente requiera, mostrando una plantilla fácil de entender.
Requisitos especiales del procedimiento:	Se debe navegar por todas las páginas para comprobar el formato común y la claridad del diseño
Resultado obtenido:	Satisfactorio

Tabla 86–PNF02

Nombre	P-NF03
Elementos de prueba	RNF-P-01
Descripción:	El sistema debe poder instalarse en Windows 10
Especificaciones de entrada:	Windows 10 como sistema operativo o una versión posterior.
Especificaciones de salida:	
Requisitos especiales del procedimiento:	Contar con un sistema operativo Windows 10.
Resultado obtenido:	Satisfactorio

Tabla 87–PNF03

Nombre	P-NF04
Elementos de prueba	RNF-P-02
Descripción:	El sistema debe poder comunicarse con el sistema de archivos de Windows (sistema operativo anfitrión) sobre el que se ha ejecutado el sistema.
Especificaciones de entrada:	Archivos de texto que el sistema descarga desde la red social Facebook.
Especificaciones de salida:	Debe guardar los de texto en un archivo en el equipo.
Requisitos especiales del procedimiento:	
Resultado obtenido:	Satisfactorio

Tabla 88–PNF04

Nombre	P-NF05
Elementos de prueba	RNF-I-03
Descripción:	El sistema debe representar la tabla en mitad de la página para que sea visible por el usuario.
Especificaciones de entrada:	Cualquier palabra del diccionario de palabras relevantes.
Especificaciones de salida:	Mostrar la tabla en mitad de la página especificada.
Requisitos especiales del procedimiento:	
Resultado obtenido:	Satisfactorio

Tabla 89–PNF05

Nombre	P-NF06
Elementos de prueba	RNF-I-04, RNF-I-05, RNF-I-06, RNF-I-07, RNF-I-08,
Descripción:	El sistema debe poder enlazar a todas las clases genéricas que estén presentes en el índice. Se deben mostrar todas estas páginas con su información correspondiente.
Especificaciones de entrada:	El usuario debe estar en la página del índice.
Especificaciones de salida:	El sistema debe mostrar la clase genérica que el usuario seleccione
Requisitos especiales del procedimiento:	
Resultado obtenido:	Satisfactorio

Tabla 90–PNF06

Nombre	P-NF07
Elementos de prueba	RNF-I-04, RNF-I-05, RNF-I-06, RNF-I-07, RNF-I-08,
Descripción:	El sistema debe poder enlazar a todas las clases genéricas que estén presentes en el índice. Se deben mostrar todas estas páginas con su información correspondiente.
Especificaciones de entrada:	El usuario debe estar en la página del índice.
Especificaciones de salida:	El sistema debe mostrar la clase genérica que el usuario seleccione
Requisitos especiales del procedimiento:	
Resultado obtenido:	Satisfactorio

Tabla 91–PNF07

Nombre	P-NF08
Elementos de prueba	RNF-I-10
Descripción:	El sistema debe ofrecer un enlace para volver a la página del índice en todas las páginas que tenga.
Especificaciones de entrada:	El usuario puede estar en cualquier página y solamente necesita pinchar en el enlace que le llevará al índice.
Especificaciones de salida:	El sistema lleva al usuario al índice.
Requisitos especiales del procedimiento:	
Resultado obtenido:	Satisfactorio

Tabla 92–PNF08

Nombre	P-NF09
Elementos de prueba	RNF-C-01
Descripción:	La comunicación de la interfaz con la base de datos se debe hacer a través de un Servlet, el cual controla las transacciones entre ambos.
Especificaciones de entrada:	Una transacción desde la interfaz hacia el Servlet que solicite datos que están almacenados en la base de datos.
Especificaciones de salida:	
Requisitos especiales del procedimiento:	La transacción se inicia por una acción del cliente en la interfaz y requiere datos que están almacenados dentro de la base de datos.
Resultado obtenido:	Satisfactorio

Tabla 93–PNF09

Nombre	P-NF10
Elementos de prueba	RNF-C-02
Descripción:	El sistema desarrollado es compatible con otros sistemas como el navegador o el ordenador anfitrión (Host).
Especificaciones de entrada:	
Especificaciones de salida:	
Requisitos especiales del procedimiento:	Sistemas compatibles con: -XAMPP v 5.6.20 o posterior - Windows 10 - Windows Explorer o Google Chrome
Resultado obtenido:	Satisfactorio

Tabla 94–PNF10

6 Micro lenguaje del segmento de la alimentación

Se ha llevado a cabo un estudio de las principales páginas dentro del sector de la alimentación online en España para encontrar las principales variables que determinan el éxito de este sector. Se han tomado como referencia asociaciones y páginas especializadas en el tema para la exploración. [10] [11]

En el estudio que se ha llevado a cabo en las distintas tiendas online se han analizado distintas características, para obtener una información completa de las páginas. Por ello se han analizado en primer lugar los bienes que se ofrecen en la página, la cantidad de términos clave que utiliza, las redes sociales en las que está y los enlaces de distinta categoría con los que cuenta la página.

Este estudio se ha llevado a cabo con distintas herramientas, documentación y recursos. Se ha usado Traffic Travis para el análisis SEO, esta herramienta es gratuita y muy completa. Se han consultado las páginas citadas anteriormente y se ha usado Excel para crear las tablas.

6.1 Lista de términos

Se ha elaborado una lista ordenando los términos por orden de importancia que se adjunta a continuación.

1. Cerezas del Bierzo:
 - a. Bierzo

- b. Vinos
 - c. Platos
 - d. Regalos gourmet
 - e. Tienda gourmet
 - f. Venta online de vino
 - g. Venta aceite
 - h. Venta de queso
 - i. Venta de jamón
 - j. Lotes
 - k. Cerezas
2. LaTiendaHero
- a. tienda online
 - b. alimentación infantil
 - c. leches infantiles
 - d. postres
 - e. frutas
 - f. cereales infantiles
3. Chuchesonline
- a. cookies
 - b. chuchesonline
 - c. chuches
 - d. chocolates
 - e. comprar chuches
 - f. online
 - g. tienda
4. FrutasZabalza
- a. verduras
 - b. frutas
 - c. magdalena
 - d. comprar
 - e. Navarra,
 - f. envasados
 - i. frutas envasados
 - ii. envasados pastas mary
 - iii. envasados frutas verduras
 - g. cestas
 - i. cestas verduras
 - ii. cestas envasados
 - iii. cestas frutas
 - iv. cestas huevos
 - h. huevos
5. Tienda Osborne
- a. Osborne
 - b. bebidas
 - c. bebidas espirituosas
 - d. alimentación
 - e. ibéricos
 - f. jamón ibérico
 - g. bellota
 - h. bodegas
 - i. vinos de jerez

- j. vinos de rioja
 - k. oportos
 - l. jamón cinco jotas
 - m. Jamón
6. Tudespensa
- a. supermercado online
 - b. super online
 - c. compra online
 - d. supermercado internet
7. Estabuenisimo
- a. regalos gourmet
 - b. comprar productos gourmet
 - c. tienda gourmet
 - d. productos delicatesen
 - e. venta de jamón
 - f. venta de vino online
 - g. venta de queso
 - h. venta de aceite
 - i. cestas de navidad
 - j. productos
8. Singluten
- a. Gluten
 - b. Alimentos sin gluten
 - c. Levadura
 - d. Celiacos
 - e. Tienda online
 - f. Galletas
 - g. Arroz
 - h. Sin gluten
 - i. Harina

Se adjunta a continuación una tabla con la repetición de estos términos para extraer cual son los más importantes.

Término	Repeticiones
Alimentación	1
Alimentación infantil	1
Alimentos sin gluten	1
Arroz	1
Bebidas	1
Bebidas espirituosas	1
Bellota	1
Bierzo	1
Bodegas	1
Celiacos	1
Cereales infantiles	1
Cerezas	1
Cestas (de navidad, frutas, etc.)	4*

Chocolates	1
Chuches	1
Compra online	1
Comprar (chuches, jamón, prod. Gourmet, etc.)	4*
Cookies y Galletas (no distingo los idiomas)	2
Envasados	1
Frutas	2
Gluten	1
Harina	1
Huevos	1
Ibéricos	1
Jamón	1*
Jamón cinco jotas	1*
Jamón ibérico	1*
Leches infantiles	1
Levadura	1
Lotes	1
Magdalena	1
Navarra	1
Online	1*
Oportos	1
Osborne	1
Platos	1
Postres	1
Productos	1
Productos delicatessen	1
Regalos gourmet	2
Sin gluten	1
Supermercado online (super online, super internet)	3*
Tienda	2
Tienda gourmet	2
Tienda online	2*
Venta	2
Venta de aceite	2
Venta de jamón	1
Venta de queso	1
Venta de vino online	2
Vinos	1*
Vinos de Jerez	1
Vinos de Rioja	1

Tabla 95- Término y repeticiones

En esta tabla hay que matizar algunos términos que han sido acompañados con un asterisco en el número de repeticiones. En el término “cestas” se ha repetido 4 veces bajo distintos términos como son cestas de frutas, cestas de navidad, de verduras, envasados y huevos. No obstante se considera conveniente englobarlas todas en un mismo término ya que en si los términos por separados no tienen mucho peso.

En el caso del término **comprar** se ha decidido aglutinar de la misma forma todos los términos que tienen alguna referencia con el término comprar, como son comprar productos gourmet, comprar jamón, comprar chuches.

En el caso del término **jamón** cabe destacar que este sí que es un término importante con mucho peso en el mercado de las ventas online. No obstante otros términos como Jamón cinco jotas y jamón ibérico, cobran mucha importancia. Por lo que se ha decidido no aglutinar estos términos en uno solo para no perder significación.

En la palabra **online** solo aparece repetida una vez no obstante no se ha de tener esto en cuenta al pie de la letra ya que la palabra online se repite con otras muchas combinaciones como son supermercado online, venta online, tienda online, etc.

Esto mismo sucede con la palabra **tienda**, ya que esta se repite a menudo en solitario y en combinación con otras. No obstante como son palabras muy importantes dentro del vocabulario no se han aglutinado todos los términos clave que poseen la palabra tienda para no perder su significado.

Los términos significativos en torno a la palabra **tienda** son tienda online y tienda gourmet, no obstante se ha podido observar que hay otros términos parecidos como tienda infantil o tienda de chuches que son significativos pero solo en páginas muy especializadas.

El término **venta** se repite mucho en las páginas que se han visitado, no obstante la particularidad de este término reside en los complementos y sustantivos que la acompañan ya que se repiten en modalidades y páginas muy distintas lo que le da mucha fuerza.

Se ha decidido no aglutinar estos términos en uno solo para no perder significación ya que algunos de ellos cobran vital importancia en las páginas especializadas en determinados productos.

El término **vino** también se repite en muchas de las tiendas online que se han visitado. No obstante este término tiene mucho peso dentro del segmento por lo que no se han aglutinado las diferentes clases de vino en una sola para no perder la posible ventaja que se puede extraer usando estos términos correctamente.

Se ha decidido agrupar en torno al término **supermercado online** otros dos términos que no aportaban ninguna significación relevante que son “super online” y “supermercado internet”. Son 2 formas distintas de expresar la misma realidad por lo que no resulta útil el mantenerlas por separado.

6.2 Los términos más significativos para el micro lenguaje.

Para llevar a cabo una lista de los términos más importantes se va a tomar en consideración el número de repeticiones del término y la importancia relativa del término en el contexto de las tiendas online. Para ello se consultará la lista del punto 4.1 como referencia.

Por ello se ha creado una nueva lista donde se almacenarán los términos de mayor a menor relevancia.

Orden	Término	Comentarios	Ponderación
1.	Tienda online		10
2.	Venta		10
3.	Comprar	Agrupando los términos de Comprar (chuches, jamón, etc.)	10
4.	Cestas	Agrupando los términos de Cestas (de navidad, frutas, etc.)	10
5.	Tienda		10
6.	Jamón		10
7.	Vinos		10
8.	Supermercado online	Agrupando los términos de super online y supermercado internet	7
9.	Venta de vino online		7
10.	Jamón ibérico		7
11.	Jamón cinco jotas		7
12.	Tienda gourmet		7
13.	Frutas		4
14.	Alimentación		4
15.	Bebidas		4
16.	Comprar online		4
17.	Vinos de Jerez		4
18.	Vinos de Rioja		4
19.	Sin gluten		4
20.	Productos		1
21.	Productos delicatessen		1
22.	Venta de aceite		1
23.	Venta de queso		1
24.	Venta de jamón		1
25.	Productos		1
26.	Productos delicatessen		1
27.	Regalos gourmet		1
28.	Ibéricos		1

29.	Bebidas espirituosas		1
30.	Osborne		1
31.	Alimentación infantil		1
32.	Bellota		1
33.	Bierzo		1
34.	Bodegas		1
35.	Cereales infantiles		1
36.	Cerezas		1
37.	Chocolates		1
38.	Chuches		1
39.	Cookies		1
40.	Envasados		1
41.	Huevos		1
42.	Gluten		1
43.	Alimentos sin gluten		1
44.	Leches infantiles		1
45.	Lotes		1
46.	Magdalena		1
47.	Arroz		1
48.	Harina		1
49.	Navarra		1
50.	Platos		1
51.	Celíacos		1
52.	Levadura		1
53.	Postres		1

Tabla 96- Término y ponderación

La ponderación se ha asignado atendiendo a los criterios de ordenación de la tabla. Lo primeros términos son los que más se han repetido y los más significativos para las consultas que se han hecho.

Por ejemplo el término compra se repite 4 veces y se sobre entiende que en este contexto es muy importante. No obstante tienda online no se ha repetido tanto el término no obstante es muy importante en este contexto.

7 Taxonomía del micro lenguaje:

Para llevar a cabo una estructuración de la información recuperada sobre los términos más populares en la red se necesita compararlos contra algún estándar de taxonomía ya existente. En este caso se ha consultado un estándar europeo de la EFSA (<http://www.efsa.europa.eu/sites/default/files/assets/804e.pdf>) sobre la clasificación de los alimentos que es muy concreto para ayudar en la realización de la estructura del lenguaje.

También se ha consultado otro estándar de clasificación de los alimentos del ministerio de sanidad de Irlanda que ayuda a completar algunos aspectos de la clasificación de los productos alimenticios (<https://www.fsai.ie/assets/0/86/204/d7af0602-b3e1-4788-802a-0684b3199e91.pdf>).

No obstante hay que enfatizar que estas dos guías son demasiado estrictas como para seguirlas al pie de la letra por lo que se opta por hacer una versión simplificada de éstas, como puede ser una guía simplificada parecida a la que se muestra en el siguiente enlace como base para establecer un estándar de codificación (<https://github.com/Open-Food/Open-Food-Standard/blob/master/datasets/Food%20Types%207%20Feb%202012.csv>).

Finalmente se ha establecido una lista de palabras relevantes que actuará como diccionario de las palabras relevantes necesarias para completar el análisis. Este análisis determinará la relevancia o no del comentario (post, mediante la medida TF.

Se puede observar a continuación como se ha conformado la base de datos con estas palabras en un documento .csv:

	A	B	C	D	E
1	Bebidas				
2	Bebidas	Refrescos			
3	Bebidas	Refrescos	Cola		
4	Bebidas	Refrescos	Cola	Pepsi	
5	Bebidas	Refrescos	Cola	Coca Cola	
5	Bebidas	Refrescos	Cola	Coca Cola light	
7	Bebidas	Refrescos	Cola	Coca Cola Zero	
3	Bebidas	Refrescos	Naranja		
3	Bebidas	Refrescos	Naranja	Fanta Naranja	
0	Bebidas	Refrescos	Naranja	Kas Naranja	
1	Bebidas	Refrescos	Naranja	Trina	
2	Bebidas	Refrescos	Limón		
3	Bebidas	Refrescos	Limón	Fanta Limon	
4	Bebidas	Refrescos	Limón	Sprite	
5	Bebidas	Refrescos	Limón	Kas	
6	Bebidas	Refrescos	Limón	Trina	
7	Bebidas	Refrescos	Limón	7up	
8	Bebidas	Refrescos	Limón	Limón & Nada	
9	Bebidas	Refrescos	De té		
0	Bebidas	Refrescos	De té	Nestea	

Ilustración 20. Visión de términos clave. -Fuente propia

8 Descripción de la empresa.

Nombre de la empresa	LULAF (Luis Lacadena Análisis de frecuencias)
Localización	Madrid
email	luislacadena@gmail.com
Sitio web	Aun no disponible.
Fundada en	Madrid 1 de enero de 2016
Administrador	Luis Lacadena Pascual

Tabla 97- Creación de la empresa

8.1 Perfil de la empresa

8.1.1 Descripción del negocio

La empresa se dedica a desarrollar un sistema capaz de ofrecer servicio gratuito limitado de recuperar la frecuencia de términos en páginas de Facebook dentro del sector de la alimentación. Esto puede ayudar enormemente a sus clientes a encontrar términos clave para sus páginas web y dirigir campañas publicitarias o simplemente conocer tendencias. Las versiones más completas del sistema serían de pago.

8.1.2 Objeto social

Diseño de sistemas informáticos para recuperar información, dando soporte online mediante una página web. En la página se ofrecen gratuitamente un servicio limitado con una serie de términos que el cliente puede buscar dentro del sector de la alimentación. Para agencias y clientes más grandes se ofrece un servicio más amplio y flexible donde se incorporarán más funcionalidades.

8.1.3 Mercado objetivo

El mercado objetivo de la empresa son empresas de publicidad, particulares, trabajadores autónomos, PYMES y asociaciones de diverso tamaño que tengan interés en conocer las tendencias dentro de páginas de Facebook dedicadas al sector de la alimentación. LULAF ofrece sus servicios para analizar la red social Facebook para encontrar términos clave e información de interés para sus clientes.

8.1.4 Clientes objetivo

Los clientes a los que se pretende llegar ocupan un amplio abanico como se ha mencionado. Ya que se pretende llegar desde los más pequeños, que serían usuarios individuales, hasta los más grandes que pueden ser asociaciones o agencias de publicidad interesados en el campo de la alimentación. Estos clientes requieren de información clara y sintetizada de Internet. Esta información suele ser difícil de encontrar debido a la gran variedad que hay. Es por ello que la empresa se centra en el campo de la alimentación para encontrar a clientes concretos que quieran estudiar este campo.

8.1.5 Competidores y nivel de competencia esperado

El mercado de las herramientas SEO y análisis de términos clave es grande. No obstante estas herramientas suelen ser más genéricas abarcando todos los campos que pueden. La propuesta es crear un sistema enfocado solo al campo de la alimentación. Ésto limita la competencia ya que no se conoce ninguna herramienta parecida enfocada al sector español. No obstante a largo plazo sí que pueden aparecer imitadores.

8.1.6 Metas alcanzadas

Hasta el momento LULAF ha desarrollado exitosamente un sistema capaz de analizar páginas de Facebook, a través de una API que proporciona esta red social. La red social podría llegar a comprar una API que ofrece mayores servicios cuando consiga más clientes.

8.1.7 Equipo de trabajo

La empresa está formada por un solo trabajador el cual posee el 100% del capital y el tiempo invertido en el proyecto. No obstante LULAF planea expandirse en servicios y personal en el futuro para llegar a más clientes y dar mejores servicios.

8.2 Descripción del proceso productivo

El ámbito de la empresa es nacional ya que se centra en páginas de alimentación en España. Se pretende alcanzar un alto grado de especialización en el mercado local para llegar a ser conocidos.

La actividad principal de la empresa es analizar las páginas suministradas por el desarrollador y otros agentes de interés para analizarlas y determinar su relevancia para su análisis posterior. Con otros interesados, me refiero a que pueden ser clientes o Stakeholders del proyecto los que aconsejen el análisis de las distintas páginas que crean de interés.

La empresa no posee actualmente proveedores directos de los cuales dependa el desarrollo continuado de su actividad. No obstante el primer paso a dar es darse a conocer y encontrar clientes que estén dispuestos a pagar por un servicio ampliado del sistema.

La razón de hacer que el servicio básico de la aplicación sea gratuito es crear un efecto llamada para posibles clientes y dar publicidad a la empresa, ya que actualmente la empresa no posee capital suficiente para llevar a cabo una campaña de publicidad en condiciones.

Una vez encontrados los clientes es importante que el jefe de la organización conozca sus necesidades y preferencias para adaptar el servicio a las necesidades emergente de éstos,

dándoles la confianza y tranquilidad de que el producto que se les va a entregar se hace en plazo y cumple las condiciones estipuladas.

9 Análisis de la situación interna de la empresa.

En este apartado se analizarán los aspectos de la compañía que influyen en el desarrollo de la comercialización de los productos [22]. Se determinan las debilidades y las fortalezas de la empresa.

9.1 Capacidades personales.

La empresa LULAF tiene actualmente un solo trabajador, ello es debido a la capacidad económica y al tipo de actividad. Por lo que éste soporta todo el reparto de tareas y responsabilidades. Este empleado es el encargado de realizar las tareas de administración a la vez que todo el trabajo técnico y de documentación.

Este trabajador tiene un nivel de formación alto, ya que ha completado todas las asignaturas de la carrera de Ingeniería Informática y ADE. Cuenta con experiencia en el desarrollo web y trabaja actualmente en una consultora multinacional.

La motivación del trabajador es muy importante ya que sobre él recae la responsabilidad de satisfacer las necesidades de los clientes y mejorar en la medida de lo posible el sistema desarrollado.

La empresa tiene gran capacidad de adaptación a las necesidades cambiantes de su cliente ya que al tener un tamaño pequeño y al ser una empresa nueva cuenta con un amplio abanico de posibilidades para adaptarse a los cambios que puedan suceder en el futuro.

9.2 Capacidades técnicas

La empresa posee un nivel tecnológico alto, debido a que es una empresa del sector informático dedicado a un producto enfocado hacia la tecnología de Internet y concretamente hacia las herramientas SEO y el análisis de la red social Facebook.

Es por ello que el trabajador debe estar en contacto con las últimas tecnologías, ésto es debido a que el ambiente exterior es cambiante y éso exige al trabajador que se forme continuamente con las nuevas herramientas que salen al mercado.

La empresa dispone de equipamiento informático suficiente para desarrollar la actividad inicial. No necesita de almacenes ni lugares físicos disponibles para una reunión por lo que el único material crítico es el ofimático.

No obstante LULAF aún no dispone de patente sobre su proyecto registrada a nivel europeo o Internacional debido a su alto coste. No obstante en el futuro es muy probable que lo haga por miedo a que surja algún competidor que copie y patente la idea.

9.3 Capacidades comerciales

En el caso de LULAF, se ofrece un sistema con una funcionalidad limitada de uso gratuito para darse a conocer y captar posibles clientes. A los nuevos clientes se les ofrecería un servicio más adaptado a ellos a cambio de un honorario que irá en función de las funcionalidades nuevas que se quieran adoptar.

Al tener solo un trabajador LULAF, depende del administrador para llevar a cabo la promoción de este servicio. Éste se puede promocionar directamente a las agencias de publicidad o a empresas que operan en el sector de la alimentación como el grupo Osborne, Tudespena.com, Ulabox, etc.

No se descarta trabajar con otras entidades como universidades y administraciones públicas para obtener información para la investigación de otros programas. Los precios se acordarán con el cliente no obstante ello irá en función de las horas dedicadas, el precio de las licencias y APIs necesarias para el sistema y los nuevos requisitos. [23]

9.4 Capacidades de gestión

LULAF se ha pensado como una PYME en el régimen de sociedad limitada para contribuir solamente por los beneficios obtenidos [24]. Con esta fórmula el trabajador evita responder con su patrimonio por las deudas que pudiese llegar a contraer la empresa.

La titularidad del inmovilizado que pueda llegar a adquirirse será de la empresa, ya que la compra de éste activo se hará a través de la empresa LULAF. Las distintas áreas administrativas de la empresa corresponden a un único trabajador que es Luis Lacadena. Por lo que él es el responsable de desarrollo, administración, comercial y financiero.

El trabajador participará por lo tanto en todas las actividades y recibirá por tanto el 100% de los beneficios al no haber un segundo socio participando en la empresa.

9.5 Capacidades financieras

La capacidad financiera indica la capacidad que tiene la empresa para invertir en el mercado con su oferta, ésta indica además cuanta capacidad se posee para invertir, para endeudarse o para responder a eventualidades no ajenas al objeto social. A través de su capacidad financiera, una empresa ejerce una buena administración de sus recursos internos y externos; asimismo, con el buen manejo de los recursos se crea una seguridad financiera que permita la suficiente liquidez para cumplir con las operaciones sin la necesidad de recurrir al alza en los costos y finalmente lograr una buena prestación de los servicios, que es el objetivo principal a alcanzar [25].

Es por ello que se explicarán los ratios de rentabilidad, no obstante éstos se desarrollarán con más intensidad en el Plan económico que se explicará en otro apartado.

Se analiza la rentabilidad económica (ROI) de LULAF, la empresa que se está analizando. Se puede conocer el rendimiento que tienen los activos. El ROI es un ratio que se obtiene con la siguiente fórmula:

$$\text{ROI} = (\text{Beneficio} - \text{Inversión}) / \text{Inversión}$$

Ilustración Fuente [75]

Aunque esta fórmula se puede expresar también como el beneficio antes de intereses e impuestos en el año t entre el activo total en el año t-1. Este ratio se ha calculado para el año 2018, esto es debido a que el negocio en los periodos anteriores se está consolidando. El resultado del ROI a finales del 2018 es 20.77%.

Otro ratio importante a analizar es el que analiza la pérdida o el beneficio con la inversión inicial de la empresa. Este ratio es la rentabilidad financiera (ROE), para el cual se usa la siguiente fórmula:

$$\text{ROE} = \frac{\text{Beneficio neto después de impuestos}}{\text{Fondos propios}}$$

Ilustración 21. Fórmula ROE. Fuente [76]

La rentabilidad se calcula para el final del año 2018 ya que, como en el caso del ROI, en los periodos anteriores (2016 y 2017), el ROE de la empresa es de 13,19%, este es mayor que cero por lo que la empresa es rentable respecto a la inversión.

El análisis del cash flow se hace en un punto aparte no obstante a modo de resumen se comenta que la empresa presupuesta mantener durante los 3 primeros años flujos de caja positivos en todo momento. El detalle de estos flujos se especifica más adelante.

El nivel de endeudamiento de la empresa se calcula dividiendo el activo circulante entre pasivo circulante. No obstante esta empresa no tiene pasivo circulante por lo que este ratio no tiene sentido hacerlo ni analizarlo.

El periodo medio de pago y cobro es instantáneo en el momento que se contrata el servicio por decisión del gestor de la empresa. Por lo que en el momento que se entrega el servicio terminado se procede a la facturación de este.

10 Análisis de la situación externa de la empresa

10.1 Análisis PEST. Variables del entorno global

Con ayuda del análisis PEST se identifican los factores del entorno general que afectan a la empresa. Este análisis suele realizarse previo a hacer un análisis DAFO. El nombre deriva de las siglas para “Político, Económico, Social y Tecnológico” [26].

Estos factores o variables de entorno no son controlables por la empresa, por lo que un cambio en ellos puede beneficiar o perjudicar a la empresa.

10.1.1 Factores políticos

La situación actual de España tras las elecciones generales de junio de 2016, es moderadamente incierta en el ámbito político ya que las principales fuerzas políticas no han llegado a un acuerdo de gobernabilidad [29].

Estas decisiones preocupan a los ciudadanos españoles y ello puede provocar un descenso en el consumo interno y la bajada de la inversión por parte de las empresas. Afectando así negativamente al crecimiento de PIB durante el año 2016.

Las previsiones del Departamento de Asuntos económicos y Europeos indican que el crecimiento económico español será positivo durante los años 2016 y 2017. Ésto es debido a que se estima que la economía crezca en torno al 2,6% durante 2016 y un 2,5% en 2017. [30]

Ello se debe a las reformas que se acataron en 2015. Estas reformas fiscales van dirigidas a reducir el impuesto de sociedades y el IRPF. Estas medidas pretenden aumentar la competitividad de las empresas españolas en el mercado internacional e interno. Aumentando así el consumo, reactivando la economía y creando más empleo.

España es un país miembro de la Unión Europea por lo que responde ante los organismos de esta organización de países. No obstante la tributación de las empresas con sede en España conlleva la obligación de pagar impuestos estatales y municipales.

Los impuestos estatales son los que se recaudan directamente por el estado. Algunos de estos impuestos son los directos como son el IRPF y el impuesto de Sociedades. Hay otros impuestos indirectos como el IVA que gravan a las empresas indirectamente.

Los impuestos municipales son recaudados por los ayuntamientos del municipio en el cual se localiza la empresa. En el caso de la empresa a la que se refiere este documento tiene sede en Madrid. Está sujeta a impuestos como el Impuesto sobre bienes inmuebles (IBI) y el Impuesto de actividades económicas (IAE).

10.1.2 Factores económicos

Como se ha mencionado anteriormente las previsiones de crecimiento económico para España en el año 2016 son positivas ya que se estima un aumento del 2.6%. El crecimiento del PIB en el año 2015 fue del 3,2%, los pronósticos para 2016 apuntan a que bajará debido al estancamiento de la demanda nacional [32].

El crecimiento del PIB en España durante el 2015 ha tenido como protagonista a la Comunidad de Madrid ya que creció un 3,4%. Solo la Comunidad Valenciana lo superó con un 3.6%. Esta media de crecimiento de España y Madrid están por encima de la Unión Europea que está en torno al 1,8% [33]. En la ilustración que se ofrece a continuación se presentan los datos.

Tasa de crecimiento anual del PIB en 2015 en términos de volumen

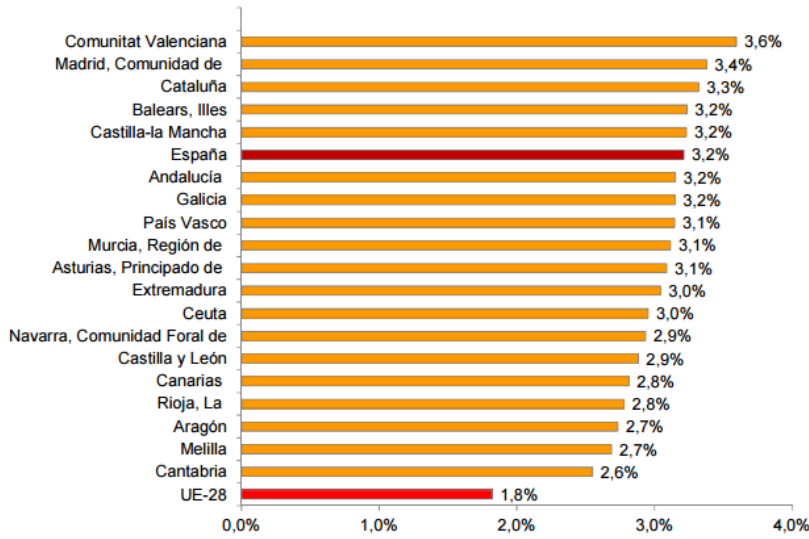


Ilustración 22. Tasa de crecimiento del PIB. Fuente [33]

El desempleo en España es un problema grave ya que las altas tasas de desempleo que se registraron en año pasado fueron del 20,9%. Esta alta tasa de paro afecta a más de 4,5 millones de personas. No obstante si se observa el gráfico inferior se observa que la tasa de paro está en una tendencia a la baja respecto a los ejercicios anteriores desde el 2013.

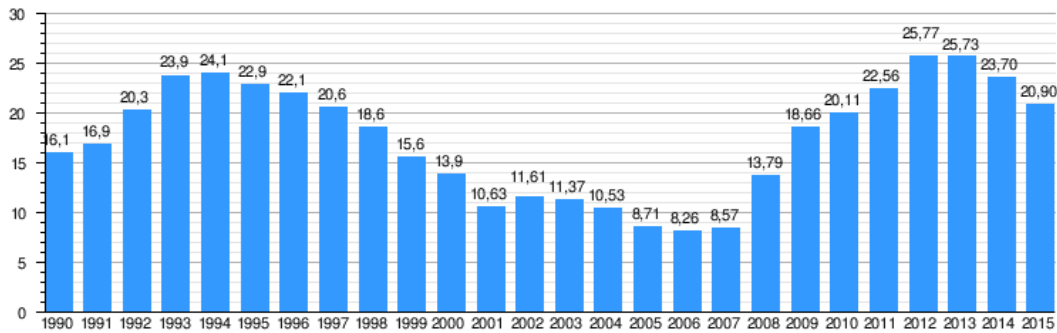


Ilustración 23. Gráfica del paro. Fuente [34]

Las previsiones de paro son favorables ya que se estima que en el próximo año siga decreciendo, llegando al 17,7% a finales del 2017.

Previsiones del paro en España para 2016 y 2017

País	1T15	2T15	3T15	4T15	2015	1T16	2T16	3T16	4T16	2016	1T17	2T17	3T17	4T17	2017
EE.UU.	5,5	5,3	5,1	5,0	5,0	5,0	4,9	4,9	4,8	4,8	4,8	4,7	4,7	4,7	4,7
ESPAÑA	23,8	22,4	21,2	21,2	21,2	21,2	20,2	19,7	19,6	19,6	19,6	18,5	18,0	17,7	17,7
UEM	11,2	11,0	10,7	10,6	10,6	10,6	10,5	10,5	10,4	10,4	10,3	10,2	10,1	10,1	10,1
PORTUGAL	13,5	11,9	11,9	11,8	11,8	18,8	11,7	11,6	11,6	11,6	11,4	11,3	11,2	11,2	11,2
R. Unido	5,6	5,6	5,5	5,5	5,5	5,4	5,4	5,3	5,3	5,3	5,3	5,3	5,3	5,2	5,2
JAPÓN	3,4	3,4	3,4	3,2	3,2	3,2	3,2	3,2	3,2	3,2	3,2	3,2	3,2	3,2	3,2
CHINA	4,1	4,0	4,1	4,2	4,2	4,2	4,2	4,3	4,3	4,3	4,2	4,2	4,3	4,3	4,3
BRASIL	5,8	6,7	7,6	7,4	7,4	7,5	8,0	8,5	8,8	8,8	8,0	8,5	9,0	9,5	9,5

Ilustración 24. Previsión del paro. Fuente [35]

La facilidad en España para abrir un nuevo negocio se puede contrastar con un indicador que es el Doing Business ofrecido por el banco mundial. En él hay 189 países. Este índice se redacta a partir de 10 variables. Se puede observar que España ocupaba el puesto 34º en 2014 y ha ascendido a la posición 33ª en el año 2015. Esta mejora supone la facilidad a inversores nacionales y extranjeros para emprender nuevas empresas en España [36].

10.1.3 Factores socioculturales

Los factores socioculturales analizan las variables como la evolución demográfica, distribución de la renta, movilidad social, cambios en el estilo de vida, actitud consumista, nivel educativo y patrones culturales [26]. Estas variables afectarán a largo plazo o de forma inmediata a los patrones de consumo y la economía del país.

La pirámide de población en España en el año 2015, demuestra que el número de hombres y mujeres en edades inferiores a 64 años es parecido. Esto incluye al grupo de jóvenes (0-15 años) y a los adultos (15-64 años). No obstante a partir de los 64 años debido a la mayor mortalidad masculina hay un menor número de hombres. Se adjunta una gráfica a continuación:

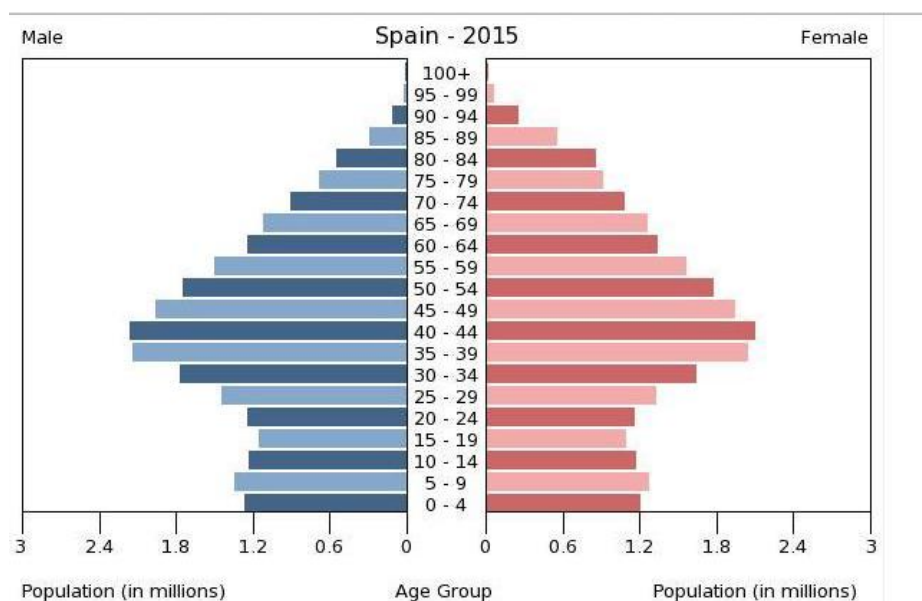


Ilustración 25. Pirámide de población. Fuente [37]

La esperanza de vida española en el año 2015 es de 83,2 años, es una de las más altas registradas por la OCDE. La media registrada por este organismo es de 79,89 años.

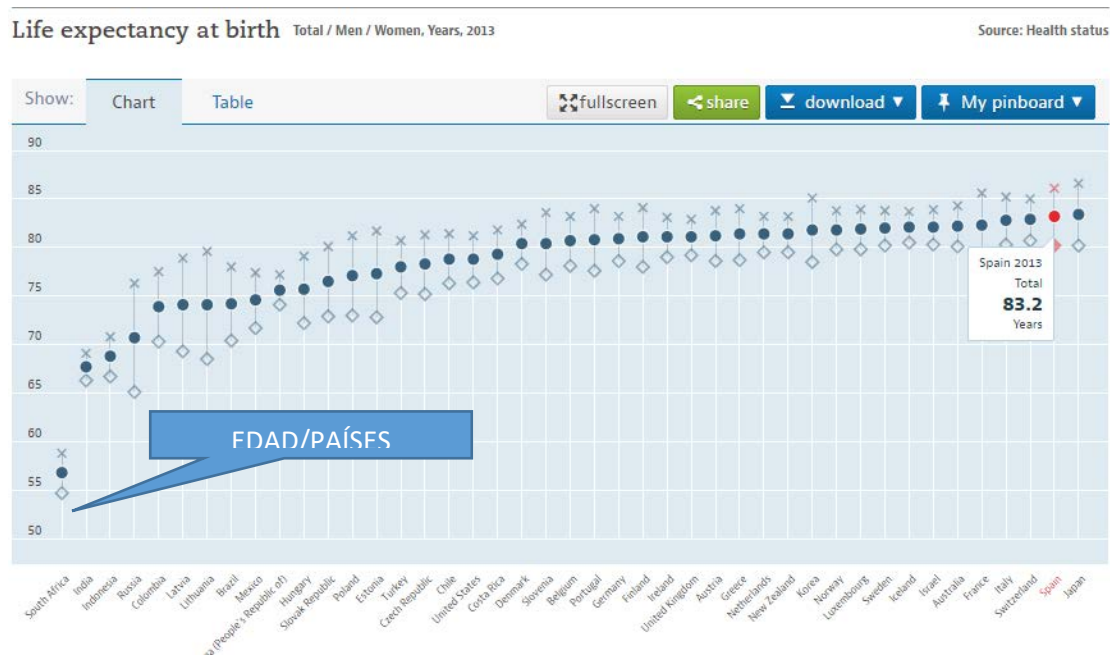


Ilustración 26. Tabla de esperanza de vida Edad/Paises. Fuente [38]

Agrandando la ilustración anterior para observar minuciosamente los resultados.

En otros puntos se ha comentado el avance de los smartphones y la tecnología de internet en la sociedad española. El 78,7% de las personas con edades comprendidas entre los 16 y los 74 años se conecta de forma regular a la red social. Este progreso sitúa a España en una posición avanzada respecto a los países colindantes [39].

El avance de esta tecnología hace que cada vez más la sociedad española se relacione de forma más estrecha con las nuevas tecnologías y los servicios que se ofrecen en la web. El gran abanico de usuarios de todas las edades, hace que se amplíe el espectro de servicios que se pueden llegar a ofertar en el futuro abriendo un próspero porvenir a las aplicaciones web.

10.1.4 Factores tecnológicos

Como se ha mencionado en el párrafo anterior un 78,7% de la población entre los 16 y los 74 años usa de forma regular las redes sociales en el año 2014. En el año 2016 se estima que ese porcentaje ha aumentado hasta el 81%, lo que demuestra una clara tendencia ascendente en los últimos años [40].

La venta de dispositivos móviles ya ha superado a la de los ordenadores y el uso de internet desde dispositivos móviles también supera al de los ordenadores por lo que se vive una transformación en el uso clásico de Internet que se venía haciendo hasta ahora [39].

Se estima que el comercio electrónico en España llegue a un volumen de 24.700 millones en 2016. Esta cifra triplicaría el resultado de 2009. Por lo que se observa la tendencia alcista del comercio electrónico. El desembarco de Amazon supuso un gran vuelco en este sentido, ya que avivó mucho la competencia del mercado [41].

La mejora de la red 4G en España y la rápida expansión de la fibra óptica hacen que cada vez sea más fácil el acceso a los canales de comercio electrónico por parte de los consumidores. Este cambio de tecnología ha cambiado los canales de compra y promoción por lo que las empresas tienen gran interés en las redes sociales.

El 54% de las PYMES están en las redes sociales, ésto es debido a su bajo coste y la diversidad de contenido y clientes que se pueden ofrecer y alcanzar en estas redes sociales de uso masivo [42].

10.2 Análisis de las 5 fuerzas de Porter. Variables del entorno específico

Las variables de entorno describen el modelo estratégico de una empresa para poder desarrollar efectivamente una estrategia de negocio. Este análisis explora las 5 fuerzas que determinan la intensidad de la competencia y la rivalidad en una industria concreta. También se puede analizar como de atractiva es esta industria en relación a oportunidades de inversión y rentabilidad [43].

Estas 5 fuerzas afectan al entorno inmediato de la organización, y afectan a la habilidad de ésta para satisfacer a sus clientes y obtener la rentabilidad esperada. Las 5 fuerzas se distribuyen en 2 subtipos. El primer subtipo son las fuerzas de competencia horizontal que son: la amenaza de productos sustitutos, amenaza de nuevos entrantes y la rivalidad entre competidores. El segundo subtipo son las fuerzas de competencia vertical que son: el poder de negociación de los proveedores, y el poder de negociación de los clientes [43].



Ilustración 27. Porter. Fuente [43]

10.2.1 Poder de negociación de los Compradores o Clientes

LULAF es una empresa que ofrece una herramienta para averiguar las frecuencias de los términos clave que aparecen en determinadas páginas de alimentación en Facebook. El uso de este sistema está dirigido a PYMES, autónomos y particulares que estén interesados en averiguar la popularidad en Facebook de los términos clave que ellos necesiten.

El poder de negociación de los clientes es alto ya que disponen de mucha información sobre este tipo de herramientas en Internet, hay muchos sistemas de posicionamiento SEO y los clientes pueden presionar en el precio debido a la información que hay en el sector.

No obstante los clientes no están organizados ya que son pequeños y muchos. Pero al poder adquirir información del sector fácilmente por vía de blogs y páginas web relacionadas con el tema, es difícil retenerlos a menos que se les ofrezca un trato personalizado.

Es por ello que LULAF ofrece un trato personalizado a los clientes adaptando la herramienta a sus necesidades. Por lo tanto LULAF se distingue de sus competidores en que ofrece un servicio orientado a un campo concreto y especializado.

En el sector de la alimentación hay muchas empresas que están en el mercado online como se ha visto en los anteriores epígrafes. Estas empresas necesitan de servicios de información y posicionamiento.

Por ello se estima que el volumen en número de ventas sea alto pero con un margen estrecho y un precio ajustado.

10.2.2 Poder de negociación de los Proveedores.

En los sectores industriales donde hay gran cantidad de proveedores, hay un poder de negociación más disipado y por lo tanto más débil. Ésto es debido a que a la empresa le resulta más fácil seguir ejerciendo su actividad normal si cambia de proveedor o si éste queda temporalmente indispuerto.

En el caso de LULAF los proveedores de aplicación y de sistemas de desarrollo y entornos de programación son bastante numerosos y existen APIs alternativas que el desarrollador puede

usar. El cambio de entorno y de API se puede hacer razonablemente rápido y a un coste bajo. Lo que no supone un perjuicio mayor para la empresa que el coste reducido para la adaptación.

En LULAF hay una amplia gama de sistemas, información y tecnologías que provienen de fuentes muy diversas. Por lo que los proveedores de estos elementos no están concentrados, ni organizados entre ellos.

Por lo que se concluye que el poder de negociación de los proveedores es bajo por lo citado anteriormente.

10.2.3 Amenaza de nuevos competidores entrantes

La amenaza de nuevos competidores mide la facilidad que tienen nuevas empresas para entrar en el negocio. Esto supondrá una amenaza para la empresa si existen pocas barreras de entrada que dificulte a los nuevos competidores entrar al mercado[49].

Según Portes hay 7 barreras de entrada que se pueden usar para crear dentro de la organización una ventaja competitiva, que son [49]:

- Economías de escala
- Diferenciación del producto
- Inversiones de capital
- Desventaja en costes independientemente de la escala
- Acceso a los canales de distribución
- Política gubernamental
- Barreras a la entrada

En el sector de las herramientas web y las herramientas SEO hay pocas barreras de entrada ya que Internet cada vez es más global y su uso está cada vez más extendido en la población española. Es por ello que cada vez surgen más proveedores de servicios en la web. Las barreras de salida son mínimas ya que desinvertir es rápido y poco complicado debido a que las licencias suelen ser mensuales y algunas anuales lo que facilita enormemente la salida.

Los competidores de esta aplicación pueden localizarse en cualquier punto de la geografía española o fuera de ella. Por lo que hay un gran abanico de posibles competidores para este tipo de empresas.

No se ha encontrado ninguna aplicación igual a la planteada, ya que las que se han encontrado son más genéricas y se centran en la figura del community Manager. No obstante el mercado de este tipo de herramientas no está saturado, aunque existen algunos competidores grandes.

El producto está diferenciado de sus competidores ya que está enfocado a un campo muy concreto que es el de la alimentación. Al centrarse en este campo se obtiene una especialización que aventaja al sistema en el conocimiento más concreto sobre el campo en cuestión.

Los costes de entrada a este mercado son bajos ya que para una aplicación de este tipo se requieren unos costes en licencias y dominios bajos y que son fácilmente recuperables una vez se ha empezado la actividad al ser de pago mensual o anual la mayoría de ellos.

Las necesidades de capital son bajas debido a que la infraestructura que se necesita comprar para poner en marcha el negocio es mínima. Esto hace que las barreras de entrada sean bajas para cualquier nuevo entrante.

No se necesita tampoco de una tienda física para ejercer la actividad lo que disminuye los costes de puesta en marcha y desarrollo de la actividad. La escalabilidad del producto a volúmenes mayores es sencilla y no requiere de una adaptación cara del motor del sistema por lo que se adapta fácilmente al uso masivo del sistema.

La proliferación de los nuevos canales de comunicación por Internet como los blogs, periódicos, revistas y redes sociales hacen que sea más fácil conocer nuevas aplicaciones y sistemas que se introduzcan en el mercado.

10.2.4 Amenaza de productos sustitutos

La amenaza de que productos similares o iguales al desarrollado provocan una presión a la baja en los precios. Estos productos satisfacen la misma necesidad, por lo tanto el abanico de productos aumenta.

Para el producto de LULAF existen varios sustitutos que pueden cubrir algunas de las necesidades. No obstante LULAF ofrece un producto más especializado que su competencia, que son las principales herramientas SEO de redes sociales del mercado. Como son:

- Bufferapp
- Hootsuite
- BuzzBundle
- Audiense

No obstante hay que recalcar que estas herramientas son más genéricas y no se especializan en ningún ámbito concreto. Por lo que no existe ninguna herramienta en concreto que sea igual a la desarrollada en este proyecto.

10.2.5 Rivalidad entre los competidores

Es el resultado de las cuatro fuerzas anteriores, ya que cuantos más competidores haya en el mercado y más rivalidad entre ellos tengan, menor será la rentabilidad del sector [49]. LULAF no posee ninguna tienda física ya que el servicio que se ofrece es online.

LULAF tampoco posee stock de mercaderías, ni almacenes lo cual la hace una empresa muy dinámica y con costes fijos más bajos. La venta de sus servicios se hace directamente a los clientes finales que pueden ser empresas o particulares.

Los competidores inmediatos de LULAF son empresas que ya están posicionadas en el mercado y son conocidas en diversos foros. Algunas de estas empresas son Bufferapp, Hootsuite, BuzzBundle y Audiense [50].

Estos competidores no están organizados entre ellos y tienen precios muy dispares entre ellos. Se adjunta una ilustración donde se han recogido todos los precios de estas páginas.

Precios de herramientas SEO			
Traffic travis al año \$	97	precio al mes	8,083333
Semrush al año\$	839,4	precio al mes	69,95
Ahora para redes sociales		precio al mes	
Bufferapp al año \$	120	precio al mes	10
Hootsuite al año €	95,88	precio al mes	7,99
BuzzBundle al año €	199	precio al mes	16,58333
Audiense al año \$	374	precio al mes	31,16667

Ilustración 28. Precios de la competencia. Fuente Elaboración propia.

No obstante el precio ronda entre los 7,99 y los 31 €, aunque no todos los servicios se pueden contratar mensualmente sino que hay algunos que solo admiten el pago de una licencia anual por el uso de la herramienta. Esto significa que los usuarios soportan un coste mayor ya que no pueden fraccionar el servicio a su medida [51].

11 Análisis DAFO

El análisis DAFO tiene la ventaja de poder ser aplicado en cualquier situación de gestión, tipo de empresa o área de negocio. Se pueden analizar empresas sin importar su tamaño o actividad [52].

Lo primero es describir la situación actual de la empresa o el departamento en cuestión, identificar las estrategias, los cambios que se están produciendo en el mercado y las capacidades y limitaciones de la empresa propia. Esta investigación servirá como base para realizar un análisis histórico, casual y proyectivo [52].

Es por ello que se analizan los factores internos y externos para identificar correctamente las debilidades, fortalezas, amenazas y oportunidades de la empresa. Estas variables se describen brevemente a continuación:

- **Amenazas:** son factores externos a la empresa que pueden poner en serio riesgo a su desempeño habitual, reduciendo su actividad productiva y por tanto su efectividad.
- **Oportunidades:** son factores externos, en los que se presentan una serie de posibilidades para la mejora de la situación de la empresa. Esta mejora de situación puede derivar en una ventaja competitiva respecto a las demás, no obstante debe ser correctamente identificada y explotada.
- **Debilidades:** son factores internos de la empresa, que provocan una situación desfavorable para la misma respecto a los competidores. Esto competidores pueden ser los actuales o los potenciales.
- **Fortalezas:** son factores internos que tiene la empresa para poder diferenciarse de sus competidores. Estos factores o recursos dan a la empresa una ventaja competitiva respecto a las demás.

11.1 Oportunidades

Las oportunidades de las que goza LULAF son:

- Herramienta especializada en el campo de la alimentación.

- Ausencia de otras herramientas centradas en la alimentación, por lo que goza de exclusividad.
- Ofertar un nuevo servicio.
- Aumento en el uso de las redes sociales a nivel nacional y global.
- Aumento de la presencia de las empresas en las redes sociales.

11.2 Fortalezas

Las fortalezas de las que goza LULAF son:

- Producto diferenciado y especializado en un ámbito.
- Costes fijos muy bajos.
- Personal cualificado.
- Empresa pequeña y ágil para adaptarse a cambios.

11.3 Amenazas

Las amenazas son:

- Nueva legislación de privacidad en la red social Facebook.
- Cambios en las tendencias de uso de las redes sociales.
- Nuevos competidores.
- Pueden surgir imitadores.
- Presión en los precios debido a la competencia del sector.
- Nuevas tecnologías de análisis de redes sociales.

11.4 Debilidades

Las debilidades son:

- Encontrar clientes.
- Dificil acceso a la financiación externa por las características de la empresa.
- Inversión en Marketing limitada por los recursos económicos de la empresa.
- Dependencia del capital humano.

12 Objetivos estratégicos

Los objetivos estratégicos son aquellas metas y estrategias planteadas por una organización para lograr determinados objetivos, y a largo plazo (entre 2 y 5 años) la posición de la organización en un mercado específico [53].

Se deben realizar determinadas acciones que permitan a la empresa cumplir con su misión, inspirándose en la visión [53]. Los objetivos deben ser claros, coherentes, alcanzables, motivadores y medibles.

- Se deben modernizar los sistemas y tecnologías de la empresa para crecer en el mercado. Esto debe hacerse proporcionando unos servicios de calidad que puedan hacer frente a los principales competidores. Esto es importante ya que la competencia en este sector exige cambios cada poco tiempo. Por lo que se debe prestar atención a los movimientos de la competencia.
- La estrategia actual de la empresa es distinguirse por su especialización en el campo de la alimentación. Esto la diferencia claramente de sus competidores, ya que por el momento es una herramienta única. Esta diferenciación debe ser explotada en el futuro sin dejar atrás el abarcar más campos para completar la herramienta.
- Para enfrentar una posible expansión de la empresa a largo plazo, la empresa tiene la opción de acudir a préstamos externos con entidades bancarias. Este apalancamiento puede mejorar la rentabilidad aprovechando el efecto del escudo fiscal [54].
- Para ampliar la cuota de mercado de la empresa, se plantea una estrategia de ampliación de los campos que abarca la herramienta pudiéndolos ampliar dentro del sector de la alimentación u otros ámbitos de interés que pueden ser sugeridos por futuros clientes.

13 Plan de marketing

El plan de marketing es un documento en el que se definen de forma estructurada los objetivos comerciales a conseguir en un periodo de tiempo determinado. Se especifican las estrategias y las acciones que se van a acometer para que sean alcanzados en un plazo determinado [55].

13.1 Plan de marketing mix

Forma parte de las estrategias del plan de marketing que definen cómo se van a conseguir los objetivos de la organización. Se deben tomar las decisiones estratégicas que engloban las 4 Ps, que son: producto, precio, distribución y comunicación [56].

Es por ello que se debe encontrar la solución a los problemas de qué servicio ofrece la empresa, el precio al que se va a ofrecer ese producto, dónde se va a vender el servicio y cómo se va a llevar a cabo la promoción del producto para darlo a conocer.

Este plan de marketing ha sido utilizado en múltiples asignaturas de la carrera por lo que ante la experiencia de uso y el conocimiento de esta herramienta, se ha preferido el Marketing mix frente a otras herramientas existentes.

13.1.1 Producto

Para analizar el producto que se ofrece, se debe tener en cuenta de que es una herramienta que ofrece un servicio a uno o varios clientes. Este servicio se consume en el momento en el que se presta al cliente.

El servicio que se ofrece, proporciona beneficios cognitivos al cliente, este es un producto potencial que va destinado a anticiparse para satisfacer necesidades demandadas por los clientes [57].

El servicio que ofrece LULAF es una página de consulta de términos clave para el sector de la alimentación. En ella los clientes pueden consultar la frecuencia de uso de los términos que se ofrecen en la página dentro del catálogo de páginas de la red social Facebook que estén especializadas en el ámbito de la alimentación que se han seleccionado.

No se patrocina ninguna marca en concreto, tan solo se proponen algunos nombres de marca por su relevancia en el mercado, ya que por cuestión de fines informativos para los clientes se han incluido para mayor completitud.

El ciclo de vida del producto se encuentra en fase de introducción, en esta fase inicial de lanzamiento los precios son altos y el volumen de ventas es bajo. Los costes de producción son altos y los beneficios son muy estrechos. Por lo que en esta etapa es muy importante invertir en publicidad u otras formas de promocionar el producto [58].

13.1.2 Precio

El precio del producto se calcula comparándolo contra la competencia y analizando los costes de mantenimiento del sistema. Estos costes también deben cubrir el coste que ha tenido el desarrollo durante los próximos 3 años.

Para crear una ventaja se ofrece un precio más bajo para atraer clientes potenciales para la empresa.

13.1.3 Distribución

La distribución del servicio se efectúa mediante el canal Internet ya que es un canal barato que ofrece muchas funcionalidades distintas, además es rápido y fiable. El acceso a este canal se vuelve cada día más popular como se ha mencionado anteriormente [59].

Muchas empresas de todos los sectores lo utilizan para vender sus productos o dar soporte a los mismos. Se puede dar un servicio fiable y de calidad sin comprometer la confidencialidad del cliente, ni los datos de la herramienta [60]. Este canal además es muy rápido y permite una comunicación bidireccional ofreciendo una comunicación más completa al cliente.

13.1.4 Promoción

Es necesario estudiar e identificar al público objetivo del producto antes de que se presente al público el servicio. El público objetivo son empresas pequeñas o PYMES y particulares, que requieran de servicios para medir la popularidad de términos clave dentro del ámbito de la alimentación o de otros si así quisieran contratar el servicio.

LULAF no va a utilizar de momento los canales habituales de promoción como son prensa, televisión o radio. Sino que debido a sus recursos limitados, plantea promocionarse mediante blogs, foros y redes sociales.

Estos canales de promoción son de difusión masiva y además son gratuitos por lo que ofrecen un gran espectro de usuarios a los que se puede llegar, ya que muchos interesados suelen estar atentos a foros sobre la materia, por lo que es fácil promocionarse en esos canales.

Se pretende usar la propia red social de Facebook creando una página de promoción dentro del perfil creado en esta página para cumplir con la normativa de la página sobre las restricciones de publicidad.

También se abrirá un perfil en Google +, para crear una página donde ofertar los servicios de la empresa. De esta forma se puede llegar a los interesados de una forma rápida y efectiva.

14 Plan de Recursos Humanos.

Con el plan de Recursos Humanos se pretende determinar y analizar aquellos elementos que están en relación con la política de los trabajadores. Estos elementos son la organización funcional, definir las capacidades, cómo se estructura la plantilla, la formación y la contratación [61].

Los recursos humanos de la empresa LULAF son fundamentales ya que esa empresa se apoya en las capacidades personales, el rendimiento del empleado y los factores humanos. Al ser una empresa con un solo empleado se realizará una descripción superficial del puesto y sus responsabilidades.

14.1 Organigrama

El organigrama de una empresa es la representación gráfica de la estructura de una empresa o cualquier otra organización, incluyen las estructuras departamentales y, en algunos casos, las personas que las dirigen, hacen un esquema sobre las relaciones jerárquicas y competenciales en vigor [62].

La estructura de la empresa por el momento es sencilla ya que cuenta solamente con un empleado. Este empleado que es el fundador de la empresa abarca todas las competencias como gestor, responsable de proyecto, responsable de recursos humanos y responsable de ventas.

Si en el futuro se contratase a otro empleado se realizaría una repartición de responsabilidades, con lo que este organigrama variaría considerablemente.

14.2 Análisis de los puestos de trabajo

La empresa cuenta con un solo empleado el cual es propietario del 100% de la empresa. Este se encarga del buen funcionamiento de esta por lo que trabaja desempeñando las funciones principales y las secundarias.

Funciones que desempeña el socio fundador:

- **Gestión administrativa:** son acciones del proceso administrativo que consisten en planear, dirigir, coordinar, organizar y controlar.

- **Dirección financiera:** se encarga de resolver los problemas de financiación y los de inversión.
- **Director comercial:** encargado de los planes de comunicación con clientes reales y clientes potenciales para promocionar los servicios de la empresa.
- **Director de desarrollo:** encargado de desarrollar y mantener las versiones del sistema para los clientes de la empresa.

14.3 Reclutamiento y selección de personal

La empresa LULAF no se ha planteado por el momento ampliar a medio plazo el número de trabajadores. Esto es debido a que aún no tiene el número de clientes necesarios para soportar otro salario. Esto implica que actualmente no se haya desarrollado ningún plan de reclutamiento.

En el caso de llevarse a cabo alguna contratación, ésta sería en el largo plazo y propuesta para el área de desarrollo, con el fin de completar más áreas de exploración de la herramienta y mejorar los procesos de clasificación y selección, contratando así a una persona experta en desarrollo de programación.

14.4 Política de retribución

La política salarial más efectiva de las empresas, concretamente la retribución determinan la relación económica entre el trabajador y la empresa. La retribución se considera como uno de varios elementos motivadores, con los que la empresa trata de implicar a sus empleados en los proyectos que ésta tiene en marcha [63].

Existen diferentes tipos de retribución y a continuación se detallan los 3 más frecuentes:

- **Retribución variable:** la retribución variable va establecida en función de los resultados alcanzados. Se debe crear un sistema de medida de consecución de los objetivos contra los que se registran estos logros. Los empleados deben conocer bien esta medida para que perciban que el sistema es claro y justo. De esta forma se consigue motivarles para alcanzar las metas propuestas.
- **Retribución fija:** suele ser la retribución base de la empresa, es decir, la retribución mínima que va a percibir el empleado. Ésta suele depender del cargo que desempeña el empleado o de las funciones y responsabilidades que este tiene asignadas.
- **Retribución en especie:** esta retribución da beneficios no monetarios a los trabajadores. En este sistema cabe destacar los pagos en especie como son los bonos para comer, transporte, bonos para gasolina, días libres, guarderías para los niños, etc. Otra forma de retribución son los beneficios de previsión extrasalarial como son planes de salud, planes de jubilación, etc. Este tipo de retribución está aumentando mucho en España durante los últimos años entre la plantilla de los trabajadores. Ya que anteriormente estaba prácticamente reservado para los directivos [64].

La política retributiva de LULAF, es de retribución variable al 100%. Ésta depende de si la empresa y el empleado consiguen sus objetivos de ventas. Esta retribución se establece en

15000 € anuales repartidos en 12 pagas. Esta retribución se basa en lo que el desarrollador cobra actualmente en PWC (PricewaterhouseCoopers) con un contrato estándar.

15 Plan Económico y Financiero.

El plan económico de la empresa tiene una previsión con un horizonte temporal de 3 a 5 años. En este caso se harán a 3 años, no obstante estas previsiones están sujetas a un elevado grado de incertidumbre por lo que no se espera que se cumplan al pie de la letra.

El análisis financiero trata de evaluar la situación económica y financiera que tiene actualmente la empresa, y proyectarla hacia el futuro de la empresa. En definitiva, enjuiciar la gestión empresarial de la unidad económica para predecir su evolución futura y poder tomar decisiones con la menor incertidumbre [65].

Con la ayuda de este plan se pronostica la liquidez futura de la empresa, para ello se utilizará una cuenta de pérdidas y ganancias, soportado con un cuadro de Estado de flujos de efectivo [65].

15.1 Plan de inversión.

La inversión inicial en cualquier negocio es fundamental para poner en marcha la actividad normal que la empresa va a desarrollar. Esta inversión viene a cubrir los gastos iniciales de licencias y adquisición de bienes de producción y/o suministros. Esta inversión debe ser acorde a los recursos de la empresa para no asfixiar la rentabilidad de la misma.

Se va a desglosar a continuación el plan de inversión diferenciando entre Activo fijo (no corriente) y Activo circulante (corriente), presupuestados para la inversión.

15.1.1 Activo fijo o no corriente.

Los activos fijos son aquellos que no sufren variaciones durante el proceso de producción, ni tampoco cambian durante el ciclo de explotación de la empresa. No obstante su duración es finita, por lo que su vida útil es cuantificable. Se muestra a continuación el activo fijo:

Activo fijo no corriente				
Descripción	Importe	Cantidad	Años	
			Amortización	Amortiz. Anual
Telefono mov	100 €	1	5	20 €
Ordenador	500	1	5	100
Silla	50	1	5	10
Mesa	30	1	5	6
Impresora	50	1	5	10
Licencias	100	1	5	20
Total	830 €	6		166

Ilustración 29. Activo fijo no corriente. Fuente propia

El total presupuestado de inversión en activo fijo o no corriente para comenzar con la actividad de la empresa se estimó en 830€. El desglose de esta cantidad se muestra en la tabla anterior, en el cual se detallan los elementos básicos para iniciar la empresa.

15.1.2 Activo circulante o corriente.

Los activos circulantes son aquellos que permanecen en la empresa un tiempo menor a un año. Estos bienes suelen usarse para el desarrollo normal de la empresa durante su actividad económica. Éste se puede monetizar dentro de los 12 meses del año.

Estos activos pagan los gastos producidos por la actividad normal de la empresa. Esto obliga a que se presupuesten estos gastos para un periodo de tiempo con el fin de que se pueda hacer frente a ellos con la ayuda de la inversión inicial y los ingresos que esta tenga al principio de su actividad.

Por las características de la empresa ésta no posee activo circulante o corriente ya que las cuentas con los clientes se saldan al momento de prestar el servicio, es por ello que no se adjunta ninguna tabla para demostrar la validez de esas cuentas.

15.2 Plan de financiación.

El plan de financiación es necesario para financiar la empresa y los activos de ésta con la aportación que hace el socio fundador. La empresa se financia a largo plazo gracias a la financiación interna que aporta el socio fundador. Por lo que carece en el momento de su creación de deudas con entidades financieras.

El socio fundador aporta para fundar la empresa 20000 €, este dinero se utilizará para financiar la empresa durante el primer año de actividad. Este dinero servirá también para financiar los activos no corrientes que necesita la empresa para desempeñar su actividad normal. A partir del primer año la empresa se financia con los ingresos que recibe de la actividad normal de la misma.

Por lo que la empresa queda financiada con un 100% de recursos propios, quedándose sin apalancar, sin tener deudas con otras entidades de crédito.

15.3 Explicación de la cuenta de Pérdidas y Ganancias

En la cuenta de pérdidas y ganancias se explican las distintas fuentes de ingresos y las de coste de la empresa. La empresa es de nueva creación por lo que la estimación de las ventas se ha

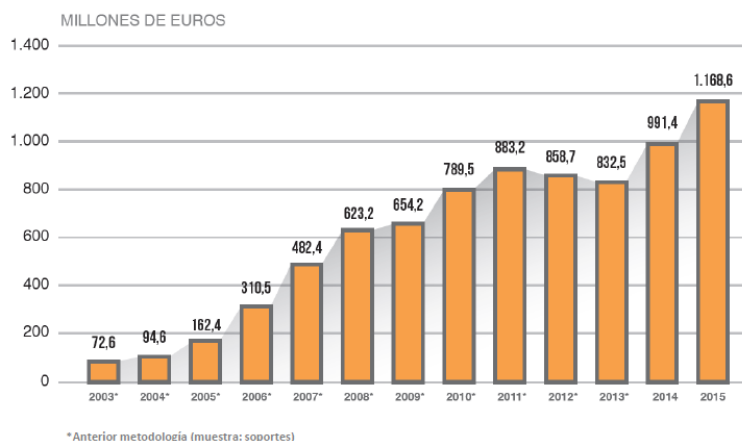
hecho basándose en los datos proporcionados por varias consultoras entre las que destacan Accenture, PWC e Infoadex.

También se ha consultado a profesionales del campo de la publicidad en medios digitales para estimar las cantidades y la facturación. Para el desarrollo de la estimación de ventas como se ha mencionado anteriormente se ha consultado una fuente en el siguiente enlace (http://www.iabspain.net/wp-content/uploads/downloads/2016/03/Estudio_Inversi%C3%B3n_Publicitaria_Medios_Digitales_2015_IAB_Spain_vReducida.pdf). Donde la agencia de marketing online IAB junto a la consultora PWC han realizado un estudio sobre la publicidad y las categorías de ésta. De este estudio se puede observar que las ventas en ese mercado siguen un patrón parecido al del PIB español, ya que si se analiza detalladamente cómo ha evolucionado el PIB de España como se muestra en la siguiente imagen:



Ilustración 30. Evolución del PIB. Fuente [72]

Y si se analizan los datos de ventas de las empresas en el informe que se ha extraído de la agencia IAB, que se puede observar a continuación:



Pregunta: ¿Cuál es la inversión total publicitaria NETA en DESKTOP –DISPLAY, SEARCH Y VÍDEO– anual?

INVERSIÓN PUBLICITARIA EN MEDIOS DIGITALES | RESULTADOS 2015

Ilustración 31. Evolución ventas online. Fuente [73]

Se observa que ambas siguen un patrón parecido durante los últimos años de la crisis económica (desde 2011 hasta 2015), pero se observa cómo durante la recuperación del PIB en 2013, la inversión en publicidad también se recupera. Cabe destacar que la recesión de la publicidad no es tan marcada como la del PIB.

Esta empresa no guarda ningún stock, ni se dedica a la fabricación de ningún bien de consumo físico. Por lo que no tiene ningún sentido hacer una la cuenta de pérdidas y ganancias por los métodos de Full y Direct Cost.

P&L	2016	2017	2018
Ingresos	23000	24311	25939,837
Aprovisionamientos	-1185	-1185	-1185
Gastos Personal	-19350	-19350	-19350
Personal	-15000	-15000	-15000
Cargas Sociales	-4350	-4350	-4350
Amortizaciones	-166	-166	-166
Beneficio Operativo	2299	3610	5238,837
Impuesto Sociedades (30%)	-689,7	-1083	-1571,6511
Resultado Neto	1609,3	2527	3667,1859

Tabla 98- P&L

Para hacer la estimación de los ingresos se ha usado la fuente [73] y la [70] para conseguir las estimaciones de crecimiento del mercado de las ventas online, ya que el crecimiento de las ventas esperado para 2016 es del 4,7% y el crecimiento para 2018 es del 6,7%. De estas dos

medidas se ha sacado la estimación para 2017 como media de la diferencia entre las dos, que sería un 5,7%.

15.4 Desarrollo del balance de situación

El balance de situación se ha desarrollado para los próximos 3 años. Para ello se ha realizado una previsión de datos de cada año hasta el 2018. Algunos de los datos se explicaran a continuación:

- Amortización: la cuenta de amortización refleja la disminución del valor de los activos con el paso del tiempo.
- Tesorería: esta cuenta representa la cantidad de dinero líquido que tiene la empresa. Ésta irá variando conforme a los diferentes ingresos y gastos que se vayan produciendo.
- Reserva legal: en las empresas españolas es obligatorio dotar el 10% de los ingresos a una reserva hasta que alcance el 20% como mínimo del capital de la empresa.
- Reserva voluntaria: son todos los beneficios que no van a la reserva legal.

15.4.1 Tabla de balance de situación

	2016	2017	2018
Activo no corriente	664	498	332
Inmovilizado intangible	100	100	100
Amortización acumulada Inmovilizado intangible	-20	-40	-60
Inmovilizado material	730	730	730
Amortización acumulada Inmovilizado material	-146	-292	-438
Activo corriente	21635	21635	29043,137
Tesorería	21635	24.721 €	29.043 €
Total	22299	25219,3	29375,137

Tabla 99- Activos

	2016	2017	2018
Pasivo neto	21609,3	24136,3	27803,4859
Capital	20000	21609,3	24136,3
Resultado	1609,3	2527	3667,1859
Reserva legal	0	160,93	413,63
Reserva voluntaria	0	1448,37	3722,67
Pasivo no corriente	0	0	0
Pasivo corriente	689,7	1083	1571,6511
Pasivos con hacienda	689,7	1083	1571,6511
Total	22299	25219,3	29375,137

Tabla 100- Pasivos

15.5 Desarrollo del Cash Flow

En este punto se presentan la previsión de las salidas y entradas de capital por las operaciones que la empresa lleva a cabo de compras y ventas. En este caso, no se ha pedido ningún préstamo a entidades financieras externas por lo que no hay pasivos con éstas ni intereses.

Se expone a continuación una tabla con las entradas y salidas de caja (dinero):

Entradas y salidas de dinero			
	2016	2017	2018
Ingreso:	43000	24311	25939,837
Capital	20000		
Ventas	23000	24311	25939,837
Salidas:	-21531	-20205,7	-20599
Aprovisionamientos	-1185	0	0
Gastos Personal	-19350	-19350	-19350
Inversiones	-830	0	0
Impuesto de sociedades	0	-689,7	-1083
Amortizaciones	-166	-166	-166
Cash flow	21469	4105,3	5340,837

Tabla 101- Cash Flow

El cash flow aumenta en cada periodo ya que se estima que el número de ventas crece a medida que pasa el tiempo. Estos datos sobre la estimación de ventas se han sacado de la fuente [70]. La salida de cash flow también aumenta a medida que lo hacen las ventas, no obstante este aumento es más lento que el dinero que generan las ventas.

15.6 Desarrollo del punto muerto

El punto muerto indica el mínimo de ventas que la empresa necesita para cubrir sus costes, estos costes pueden ser fijos o variables. Este punto se podría entender como un punto de inflexión a partir del cual la empresa empieza a obtener beneficios.

El punto muerto puede ser expresado de diferentes maneras:

- Unidades monetarias: se hace calculando el precio al que el beneficio de la empresa es cero.
- Unidades físicas: el número de unidades del producto o número de unidades respecto a los servicios que hay que prestar.

En LULAF, que es la empresa de la que se está haciendo el estudio, hay unos costes anuales (sumando costes fijos) de:

	2016	2017	2018
Salidas:	-21531	-20205,7	-20599
Aprovisionamientos	-1185	0	0
Gastos Personal	-19350	-19350	-19350
Inversiones	-830	0	0
Impuesto de sociedades	0	-689,7	-1083
Amortizaciones	-166	-166	-166

Tabla 102- Salidas de efectivo

Se procede a calcular el punto muerto de 2016 para tener una referencia aproximada del umbral de ventas de cada uno de los productos por separado. Serían:

Términos incluidos	2725,087500	aproximado	2726	Unidades
Nuevos ámbitos	9,478565217	aproximado	10	Unidades

Tabla 103- Punto Muerto

Estas cifras se corresponden a l caso en el que solo se vendiese uno de los dos productos (el número de unidades del otro producto sería 0). Por lo que se ve que el número de proyectos para investigar nuevos ámbitos sería 10 proyectos, con un coste muy superior (2300€/proyecto) al de añadir términos (se facturaría a 8€/término). Mientras que añadir términos necesita una cifra muy superior con 2726 unidades o términos incluidos para llegar a cubrir los costes de la empresa.

15.7 Análisis financiero

En el siguiente punto se calcula la Tasa interna de retorno (TIR) que mide la rentabilidad de la inversión realizada. También se calcula el valor actual neto (VAN), que mide los flujos de caja actualizados que la empresa va a generar durante su actividad.

La tasa de interés k es 0.085 ya que es la tasa de interés que aplica el banco central para sus préstamos [74].

	l. inicial	2016	2017	2018	n años	k (tasa de interés)	
Beneficios	-20000	1609,3	2527	3667,1859	3	0,085	%

Tabla 104 -VAN

15.7.1 Valor actual neto (VAN)

El valor actual neto VAN = 6500,87€

Este resultado se ha calculado usando los flujos de la empresa, estos flujos son actualizados a una tasa de 0,085% que es la tasa que aplica el bono español a 5 años. El valor del VAN demuestra que la inversión en esta empresa es rentable ya que da más que 0.

15.7.2 Tasa interna de retorno (TIR)

La tasa interna de retorno de la empresa es TIR= 24%.

Esta tasa se ha calculado usando los flujos de caja y teniendo en consideración la inversión inicial. La TIR es la tasa de rentabilidad que hace que el VAN sea igual a cero. En este caso la TIR es mayor a la tasa de interés o coste del capital k, por lo que se deduce que la inversión es rentable. El periodo que se ha usado es 10 años para que se observe que es posible recuperar la inversión inicial de 20.000€.

16 Gestión del proyecto

En este capítulo se detalla la planificación que se ha llevado a cabo para la consecución del proyecto.

16.1 Planificación del proyecto

A continuación se detallan las fases que se han seguido para llevar a cabo el Trabajo de Fin de Grado. Se ha seguido como base metodológica SCRUM, esta metodología no se ha seguido al pie de la letra. No obstante ayuda en el desarrollo y la planificación del sistema.

La planificación se realiza teniendo en cuenta que el número de créditos que corresponden al Trabajo de Fin de Grado son 24 créditos ECTS. Como indica la normativa del proyecto cada crédito ECTS corresponde a unas 25 horas de trabajo. Por lo que la dedicación al Trabajo de Fin de Grado oscilaría entre 600 horas de trabajo por parte del alumno.

La estimación de las horas de trabajo se lleva a cabo en función de la experiencia del alumno a lo largo de su etapa universitaria y fuera de ella. Para hacer la planificación se usará la herramienta Draw.io. Esta herramienta es muy sencilla y completa, ya que al ser un proyecto sencillo no se necesitan herramientas más complejas y profesionales.

La planificación del proyecto ha pasado por diferentes fases. Durante toda la vida del proyecto se irá redactando la documentación necesaria para reflejarla en la memoria del Trabajo de Fin de Grado.

Se presentan a continuación dos ilustraciones de cómo se ha llevado a cabo la planificación del proyecto. La primera incluye las fechas de ejecución de las tareas. En la segunda ilustración se muestra un diagrama Gantt de forma gráfica con las tareas.

	Task Name	Duration	Start	ETA
1	Introducción	10 días	01.01.16	11.01.16
2	Estado del arte	25 días	12.6.16	5.02.16
3	Análisis de la solución	36 días	06.02.16	13.03.16
4	Diseño	51 días	14.03.16	03.05.16
5	Especificación técnica del plan de pruebas	18 días	04.05.16	21.05.16
6	Micro lenguaje del segmento de la alimentación	16 días	22.05.16	06.06.16
7	Taxonomía del micro lenguaje	8 días	07.06.16	14.06.16
8	Descripción de la empresa	4 días	15.06.16	18.06.16
9	Análisis de la situación interna de la empresa	8 días	19.06.16	25.06.16
10	Análisis de la situación externa de la empresa	6 días	26.06.16	01.07.16
11	Análisis DAFO	3 días	02.07.16	04.07.16
12	Objetivos estratégicos	11 días	05.07.16	15.07.16
13	Plan de marketing	7 días	16.07.12	22.07.16
14	Plan de recursos humanos	8 días	23.07.16	31.07.16
15	Plan económico financiero	15 días	01.09.16	14.09.16
16	Gestión del proyecto	3 días	15.09.16	17.09.16
17	Conclusiones y trabajo futuro	6 days	18.09.16	23.09.16
18	Memoria		01.01.16	23.09.16

Ilustración 32. Planificación tiempos. Fuente propia

Se pueden comprobar las fechas de inicio, así como los días necesario para su ejecución y la ETA (Estimated time of Arrival) que se traduce en tiempo de entrega estimado.

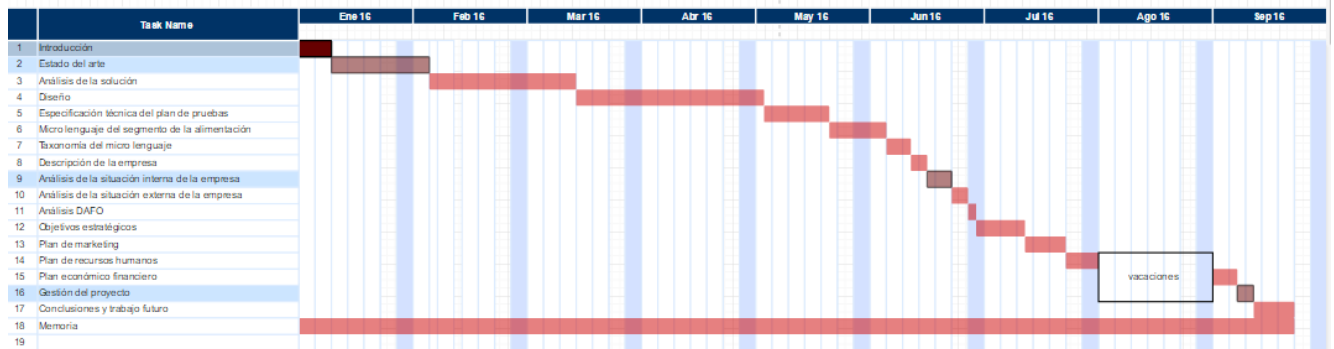


Ilustración 33. Planificación gráfica. Fuente propia

En la anterior ilustración se adjunta la planificación que se ha llevado del proyecto a lo largo de los 9 meses que ha durado. Esta planificación se ha hecho con el fin de poder compaginar estudios y proyecto. También se ha tenido que planificar el desarrollador con un trabajo en la consultora PWC.

El proyecto ha durado 631 días, esta cifra tan alta viene dada al ser un proyecto mixto y al haber tenido que compaginar estudios y trabajo con el proyecto. A una media de una hora al día se consigue sobrepasar la barrera de las 600 horas que se estima que debe durar el proyecto.

17. Conclusiones y trabajo futuro

17.1 Conclusiones

Este trabajo ha analizado algunos de los problemas que actualmente tiene la web estática a la hora de ponderar el contenido y la calidad de las páginas de reciente creación frente a las páginas más antiguas. Más concretamente ayudar a discernir la calidad del contenido de las páginas independientemente de la antigüedad de las mismas.

Algunos de estos problemas pueden llegar a ser solucionados en el futuro gracias a la tecnología de “Big data” en el que los contenidos de la página serán la llave hacia el éxito en la escalada del posicionamiento de las páginas.

Las redes sociales juegan un papel muy importante en la generación de contenido en Internet, es por eso que este trabajo se ha centrado en Facebook, que es la principal red social en el mundo, en España es la segunda con más usuarios después de WhatsApp.

Esto significa que un gran número de los internautas españoles usan Facebook habitualmente, para comunicarse, informarse y cada vez más, conocer nuevas marcas y productos. Un reciente estudio muestra que un 23% de los usuarios siguen a una marca y un 19% habla de productos en ellas como demuestra el estudio de “We are social”, una compañía de Singapur que realiza esta clase de estudios.

Ello abre un mercado para las herramientas que exploran las redes sociales para promocionar sus productos y hacer marketing online. El estudio mencionado anteriormente también indica que el 65% de los usuarios que han comprado por internet se sintieron influidos por las redes sociales a la hora de hacer su compra.

Muchos de estos usuarios siguen a las marcas para informarse sobre promociones y descuentos de las mismas. La principal red social que ejerce una influencia sobre estos consumidores es Facebook, eso me lleva a pensar que en el futuro mucha de la publicidad de las distintas marcas se focalizará en promocionarse en las redes sociales.

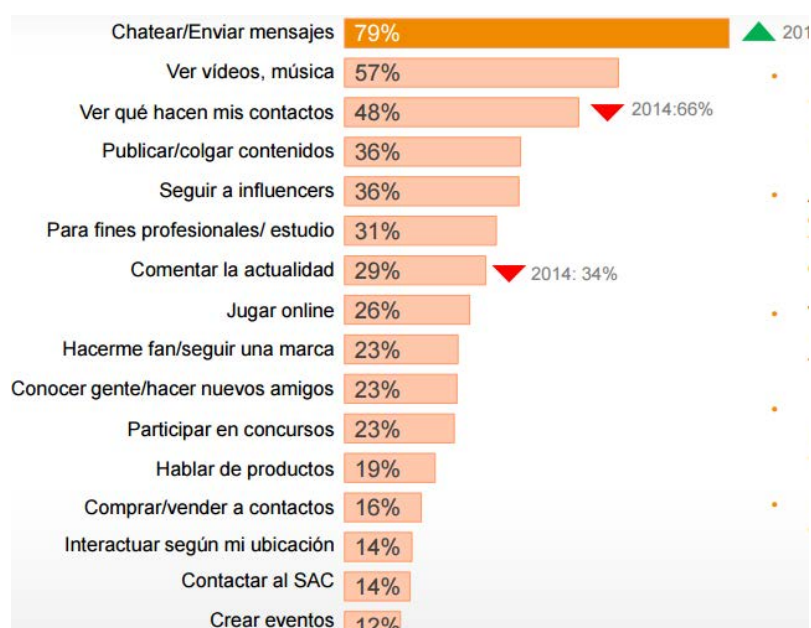


Ilustración 34. Uso redes sociales. Fuente [78]

El sector de la alimentación ocupa el quinto lugar en los sectores que más se siguen en las redes sociales, con un 25% de la audiencia. Esta transformación llega de la mano de la rápida extensión de los medios de comunicación digitales como ya se comentó en un apartado anterior. Los móviles y las tablets han abierto un nuevo horizonte de conectividad y servicios a un nuevo mercado.

El prototipo presentado analiza las redes sociales en busca de los patrones que los consumidores siguen al navegar, comentar y enriquecer las redes sociales. La investigación supone un avance hacia las preferencias de los usuarios que utilizan las redes sociales. El mercado de los productos online está en crecimiento y las expectativas son que sigan creciendo durante los próximos periodos.

La información que se recoge de las páginas de alimentación puede ayudar a conocer hábitos y patrones de consumo reconociendo el comportamiento en relación con la estacionalidad. Estos

datos pueden ayudar a empresas de publicidad y particulares a recabar información sobre los términos clave que los usuarios publican a través de sus comentarios en las redes sociales.

Hay que mencionar que el desarrollo del proyecto ha sido lento debido a que he cursado asignaturas a la vez que realizaba el proyecto. Inmediatamente después empecé a trabajar en una consultora (PWC), es por ello que se han adquirido también competencias distintas mientras estaba estudiando y luego trabajando que me han ayudado a desarrollar el proyecto y a incorporar nuevos conocimientos.

A una media de una hora al día de trabajo se han dedicado un total de 631 horas de trabajo, esto son 31 horas más de lo que se estima en las normas de la universidad que debe durar un proyecto conjunto de 24 créditos.

La creación de la empresa no es muy costosa, ya que un solo socio fundador puede hacer frente al primer año de actividad de la empresa, con un desembolso inicial de 20000€. Esta cantidad es baja si se compara con empresas de otros sectores, ya que al ser una empresa de servicios no tiene stock y el mayor capital de la empresa es el capital humano.

La empresa obtiene beneficios desde el primer año de su formación lo cual la convierte en un negocio atractivo ya que su TIR es del 26%.

17.2 Evaluación del trabajo realizado

De este proyecto he aprendido lo importante que es estudiar bien el producto antes de ponerse a hacer cualquier tarea de programación o cálculo contable. Este estudio inicial del producto es importantísimo para no dar vueltas, caer en errores innecesarios y tener retrasos en el desarrollo del producto.

El desarrollo de este proyecto me ha motivado la mejora del aprendizaje de mi conocimiento sobre los distintos procesos de desarrollo. Estos procesos de desarrollo a lo largo de la vida del software los he aprendido a lo largo de la carrera y también en mi puesto de trabajo.

He adquirido también conocimientos de nuevas tecnologías como es XAMPP y WEKA. Éstas me han ayudado mucho durante el desarrollo y para comprobar mis resultados del análisis. Esta experiencia me ha servido para reforzar también los conocimientos adquiridos en la carrera de las asignaturas de programación web para poder desarrollar el sistema del que trata en parte este proyecto.

He aprendido mucho del sector de la alimentación para la empresa que tengo montada con mis socios. Más concretamente, el desarrollo de este proyecto me ha venido bien para conocer en profundidad este segmento del mundo del comercio electrónico.

Me ha parecido difícil el dimensionar todo el sistema, pensar sus funcionalidades y escoger todas las herramientas, ya que me doy cuenta de que las funcionalidades importantes son fáciles de detectar, no obstante otras que son complementarias son más difíciles de detectar. No obstante el uso de las metodologías ágiles me ha ayudado a evitar retrasos en el desarrollo de la aplicación.

El sector de la venta de productos alimenticios en la Web está en pleno crecimiento, este crecimiento sostenido e incremental tiene la tendencia de continuar en el futuro según los pronósticos que se han consultado de la empresa Nielsen. Por lo que ante esta situación de crecimiento la creación de una empresa dedicada a este sector es viable.

La empresa creada para este proyecto tiene unas cuentas solventes y es fácil crear una pequeña sociedad con un capital reducido de 20000€ con la que se puede operar en el mercado de la venta online.

17.3 Trabajos Futuros

En este apartado se plantean algunos puntos que pueden ser desarrollados en el futuro con respecto al trabajo que ya he hecho.

Se plantea la posibilidad de que el usuario pueda descargar informes en archivos de tipo PDF para que pueda consultarlos más adelante y sea más cómodo para ellos usar la herramienta. También se incluiría un apartado donde los clientes puedan aconsejar páginas de Facebook que se puedan analizar en el futuro.

También se plantea que el algoritmo pueda analizar los “likes” o “me gusta” de las páginas y los comentarios que analiza para tener en cuenta la popularidad del término que se está analizando.

Esta última idea viene de que uno de los caminos que puede tomar el mundo del SEO y de la publicidad online, es que los términos clave van a perder un poco de su fuerza frente al contenido de las páginas.

Esto sucede por la aparición del “Big data” que ha removido los cimientos del posicionamiento de las páginas web al poder cruzar contenido que no sea exclusivamente texto. Para ello se utilizan las bases de datos no relacionales que son capaces de almacenar todo tipo de contenido digital.

El contenido digital de las bases de datos no relacionales puede ser rápidamente analizado por potentes algoritmos que ofrezcan información relevante a los usuarios que no solo está guardado en forma de texto.

Es por ello que las redes sociales jugarán un papel muy importante, ya que ofrecen y también crean mucho contenido. Este contenido puede resultar valioso para las empresas y merece la pena analizarlo, en mi humilde opinión.

18.English abstract

Summary

This project aims to create a company in order to develop a tool capable of tracking key terms in the social network Facebook. The tool developed by the company is focused on the online food market. Most SEO tools do not seek information on the pages of social networks in a given field. This tool searches for the relevant terms in the relevant pages chosen on the online food market (sector). These pages have been analyzed and the relevant posts have been extracted to be analyzed in order to return to the customer the frequency of use over a period of time. This tool can be used by individuals and/or enterprises to choose the relevant terms for their web pages and the seasonality that they have, thanks to the web service that this company offers to display the results graphically.

18.1 Introduction

The aim of this work is to develop a tool that is able to automate the process of search terms and their frequency of appearance in the pages of Facebook who engage in business and / or sale of food products online.

The growing importance of the Internet for e-commerce, and specifically of social networks in promoting products, makes necessary the development of tools that monitor content publishing to improve its dissemination.

The nature of search engines and resources is primarily textual in nature, so the terminology used in them acquires special relevance. The terminology is, in this case is focused in the food sector.

We have selected a number of Facebook pages in the food sector. The reason for choosing Facebook is that it is the social network with more users, as it has 1550 million. Other social networks like Youtube has one billion, 350 million Twitter and Instagram 400 million [44].

Below is a brief state of the art website optimization, this will help you understand the problem that arises solve.

18.1.1 The approach of the issue.

The explosion of Internet use since the early nineties has caused online content has skyrocketed since. The growth of Internet content grows exponentially.

This exponential growth requires specialized elements work as generic tools satisfactorily meet the general fields of work. But with the growing volume of information on the web more powerful tools are needed.

However there is currently no concrete answer to each of the different segments of the Web. Specifically for the growing market of online sales, which is booming. This online shopping market also requires specialized tools.

Due to growth in the volume of information has become more difficult to find useful content. Therefore indexing the Web was introduced, in order to label and sort the contents as if the index of a book concerned.

This is the main idea of search engines emerged. The task of these is through the different pages to gather all the information tracked and incorporate it to indexes that maintain all this information.

It is noteworthy that these search engines have algorithms that help them deliver results to users searching for content on the web. But the key also lies in how these contents are sorted by the relevance they have for the customer.

This relevance is usually translated in the popularity of a particular page. This popularity is assigned by the management of different websites. This popularity is calculated based on the number of links that receives the page and popularity have these, which corresponds to the external positioning.

Web positioning is key to the websites of companies, as to the amount of content that is in the Web, users will read only the first results others the last ones will be ignored by most users.

This is why SEO techniques and tools help new and existing web pages to improve the structure, visibility and quality of its content to move up in the ranking of web positioning.

Currently on the market of SEO tools you can find many tools that help improve the positioning of the pages referring to different parameters and properties of the consulted page.

Although these tools provide much information on the consulted pages such as your "page rank (still not provided by Google)", "domain authority", "Alexa Rank", "key terms", number of inbound links (backlinks), presence on social networks and more. These tools are generic and do not focus on any particular segment.

However, some of these tools do not include the market leading social network is Facebook. Since this page requires the use of an API own payment for visits to private profiles due to the privacy policy on this page. Allowing access only to the pages of common interest and businesses.

Although there is also a form of internal positioning it depends heavily on distribution and how information is presented on the page. In order to make this a more responsive and with clearer content that can be easily read by search engines page.

However this system has a distinct disadvantage because it treats the web as a static element. So websites that are important are now even more in the future. So that new sites it more difficult to reach the old because in classification are much less popular.

This new website may contain quality information that offers to the users of the web. Despite not being more popular than the ancient sites is the barrier that only a few users will come to know. Due to this static view of the web, new pages are at a distinct disadvantage with the old.

18.1.2 Goals

After detailing the problem on the various existing SEO tools and disadvantages of the static view of the web it proceeds to present a scan tool on Facebook. This tool allows the user to explore the frequency of occurrence within the various pages of the food sector.

This tool aims to be an alternative to other tools current SEO, since being free and specific use for a particular area, which in this case is food, can serve the user to analyze a key term over time. Although currently only the information relating to 2015 is analyzed.

It has begun with a more specific domain such as food in order to create the first version of the tool. However it could easily be adapted to other domains if required.

It is intended therefore analyze the main market SEO tools to observe the behaviour of these. This will be analyzed the shortcomings and advantages thereof against the proposed application.

It is therefore necessary to analyze the pages devoted to the sector of online sales of food products to which this tool is aimed as they are an important and very interested in this type of tool.

The aim to achieve is a tool that is able to help users improve their SEO positioning. So as explained above for the newly created pages is difficult to compete against older pages.

18.1.3 Memory contents

This section details what is explained in each of the chapters and annexes to the document.

In the first chapter named "Introduction" provides an overview of the project and the objectives, and other content in the document.

The second chapter is "the state of the art", which explains the current context of the theme of which is to be treated. In it the Pagerank algorithm used by Google, other tools to analyze the popularity of pages, technologies that have been used for development, brief explanation of techniques Google Hacking, a brief description of the market for online food sales is explained best rated shops and 2015.

The third chapter is "Analysis of the solution" in which a definition of such as system requirements of this and the identification of the subsystems in the application is made.

The fourth chapter is called "Design", in which the system architecture is detailed, the user interface has been used to identify and specify the classes that have been used, physical design of databases and as has been done initial data migration.

The fifth chapter is "technical specification test plan, which tests to check the correct operation of the system are included.

The sixth chapter is called "Micro language of food segment ", which analyzed the relevant terms in the field of sale of food products on the Internet.

The seventh chapter is called "Taxonomy of micro language", which carried out a study of the different classifications of foods based on European standards.

The eighth chapter is the "description of the company", in which the profile of the company and the productive process that is detailed.

The ninth chapter is the "Analysis of the internal situation of the company," where personal business skills, technical, financial and management company created specified.

The tenth chapter is the "Analysis of the external situation of the company", which has made a PEST analysis to analyze some of the global environment variables. There has also been the analysis of Porter's 5 forces to study the specific environment.

The eleventh chapter is the "SWOT analysis" in which opportunities, strengths, threats and weaknesses of the company that was created in this project are analyzed.

In twelfth chapter are the "strategic objectives", in which the defined goals and strategies to achieve the long term objectives of the company.

The thirteenth chapter is the "Marketing Plan", in which product promotion plans and the company in the present and in the future are discussed.

The fourteenth chapter is the "Human resources plan", where the organization chart presents the analysis of the jobs, the recruitment and remuneration policy.

The fifteenth chapter is the "Economic and Financial Plan" where investment plans, financing detailing, cash flows, the deadlock, the profit and loss account and other data of interest to economically analyze the viability of the company.

Sixteenth chapter is the "Project Management", in which planning has been estimated and followed for this project is detailed.

The seventeenth chapter are the "Conclusions and Future Work", which briefly summarizes the work that has been done and recommendations for further development in the future are given.

The eighteenth chapter is the "English abstract" which briefly summarizes some sections of the report in English.

In the final part there are several annexes where the resulting data to help understand the document memory is full. Special terms and acronyms as well as the sources cited are also added.

18.2 Company description

Company name	LULAF (Luis Lacadena Análisis de frecuencias)
Place	Madrid
email	luislacadena@gmail.com
Web site	Not aviable.
Founded in	Madrid 1 st January 2016
Administrator	Luis Lacadena Pascual

Tabla 105- Company description

18.2.1 Company profile

18.2.1.1 Business overview

The company is dedicated to develop a system capable of offering limited free service to recover the frequency of terms in Facebook pages within the food sector. This can greatly help their customers find key terms for your web pages and direct advertising campaigns or simply knowing trends. The most complete versions of the system would require a payment.

18.2.1.2 Company purpose

Design of computer systems to retrieve information, providing online support through a web page. On page is offered free with a limited number of terms that the client can search within the food service sector. For agencies and larger customers a broader and more flexible service which is offered functionalities will be added.

18.2.1.3 Target market

The purpose of the enterprise market are advertising companies, individuals, self-employed, SMEs and associations of various sizes that are interested in trends within Facebook pages dedicated to the food sector. LULAF offers its services to analyze the social network Facebook to find key terms and information of interest to their customers.

18.2.1.4 Target customers

Customers who aims to reach a wide range occupy as mentioned. Since it is intended to reach from the smallest, which would individual users to the largest that can be associations or advertising agencies interested in the field of food. These customers require clear and synthesized Internet information. This information is usually hard to find due to the great variety. That is why the company focuses on the field of food to meet specific customers who want to study this field.

18.2.1.5 Competitors and competition level expected

The market of SEO tools and analysis of key terms is wide. However these tools are often more generic covering all fields they can. The proposal is to create a system focused only to the field of food. This limits competition and that no similar tool focused on the Spanish sector is known. However long term imitators can appear.

18.2.1.6 Goals achieved

So far LULAF has successfully developed a system capable of analyzing Facebook pages through an API that provides this social network. The social network could get you to buy an API that provides more services when you get more customers.

18.2.1.7 Work team

The company is formed by a single worker who owns 100% of capital and time invested in the project. However LULAF plans to expand services and personnel in the future to reach more customers and provide better services.

18.3 Strategic objectives

The strategic goals are those goals and strategies set by an organization to achieve certain goals, and long term (between 2 and 5 years) the position of the organization in a specific market.

They must perform certain actions that enable the company to fulfill its mission, inspired by the vision [53]. The objectives should be clear, coherent, achievable, measurable and motivating.

- Systems must be modernized and enterprise technologies to grow in the market. This should be done by providing quality services that can cope with major competitors. This is important because competition in this sector requires changes every little while. As should pay attention to the movements of the competition.
- The current strategy of the company is distinguished by its specialization in the field of food. This differs clearly from its competitors, because at the moment is a single tool. This differentiation should be exploited in the future without leaving behind the cover more fields to complete the tool.
- For a possible expansion of the company in the long term, the company has the option of going to foreign loans with banks. This leverage can enhance returns by taking advantage of the tax shield effect [54].
- To expand the market share of the company, a strategy of expanding the fields covered by the tool opening the option expand in the food sector or other areas of interest that may be suggested by future customers it arises.

18.4 Marketing plan

The marketing plan is a document that defines a structured way to achieve business objectives over a period of time. Strategies and actions to be undertaken to be achieved within a certain period specified [55].

18.4.1 Marketing mix plan

It is part of the marketing plan strategies that define how they will achieve the objectives of the organization. They must make strategic decisions that include the 4 Ps, which are: product, price, place and promotion [56].

That is why we must find the solution to the problems which service the company offers, the price at which they will offer that product, where it will sell the service and how to do is to bring product promotion for make it known.

This marketing plan has been used in multiple subjects in my degree so to use the experience and knowledge of this tool has been preferred over other marketing mix existing tools.

18.4.1.1 Product

To analyze the product offered, you should be taken into account that it is a tool that provides a service to one or more clients. This service is consumed at the time it is provided to the customer.

The service offered, provides cognitive benefits to the customer, this is a potential product that is intended to anticipate to meet needs demanded by customers [57].

The service offered is a page LULAF query key terms for the food sector. Here customers can check the frequency of use of the terms offered in the catalog page within the pages of the social network Facebook that are specialized in the field of food that have been selected.

No particular brand is promoted, only some brand names are proposed by their relevance in the market, as a matter of information purposes for clients are included for completeness reasons.

The product life-cycle management is being introduced, in this initial phase of launch prices are high and turnover is low. Production costs are high and benefits are very narrow. So at this stage it is very important to invest in advertising or other ways to promote the product [58].

18.4.1.2 Price

The product price is calculated by comparing it against the competition and analyzing the costs of system maintenance. These costs must also cover the cost has been the development over the next 3 years.

To create an advantage for a lower price is offered to attract potential customers for the company.

18.4.1.3 Place

Service delivery is made through the Internet channel as it is a cheap channel that offers many different features, it is also fast and reliable. Access to this channel becomes increasingly popular as mentioned previously [59].

Many companies from all sectors use it to sell their products or support them. It can give a reliable and quality service without compromising client confidentiality, or the tool data [60]. This channel is also very fast and enables bidirectional communication offering a more complete customer communication.

18.4.1.4 Promotion

We need to study and identify the target audience of the product before it is presented to the public service. The target audience are small businesses or SMBs and individuals who require services to measure the popularity of key terms in the field of food or other if they wanted to hire the service.

LULAF will not use the usual time promotional channels as are newspaper, television or radio. But due to limited resources, poses promoted through blogs, forums and social networks.

These promotional channels are mass media and they are free so they offer a wide spectrum of users that can reach as many stakeholders tend to be aware of forums on the subject, making it easy to advertise on these channels.

It is intended to use the social network Facebook itself creating a promotional page in the profile created on this page to comply with the regulations of the page on advertising restrictions.

It will also open a profile on Google +, to create a page where offer the services of the company. This way you can reach those interested in a quick and effective way.

18.5 Conclusions and future work

18.5.1 Conclusions

This project has talked about some of the problems of the problems of the static web and the disadvantages posed for newly created web pages. However the future promises new solutions thanks to the new phenomenon the "Big data" where contents of the web pages will be the key to positioning success on browsers.

Social networks play a n important role in generating content on the Internet, thats why this project has focused on Facebook, which is the leading social network in the world, in Spain is the second one after Whatsapp.

This means that a lare number of Spanish Internet users use Facebook regularly to communicate, get informed and incresingly get in contact with new brands and products. A recent study shos that 23 % of users of users follow a brand and a 19 % speak about the products of that brand.

Which opens a market for tools that explore social networks to promote their products and make online marketing. The above study also indicates that 65% of users who have purchased goods through the Web were influenced by social networks in their decision.

Many of these customers follow these brands on social networks to be informed about promotions and discounts. The main social network that influence these consumers is Facebook, that leads me to think that in the future a lot of advertising from the different brands will focus on promoting themselves in social networks.

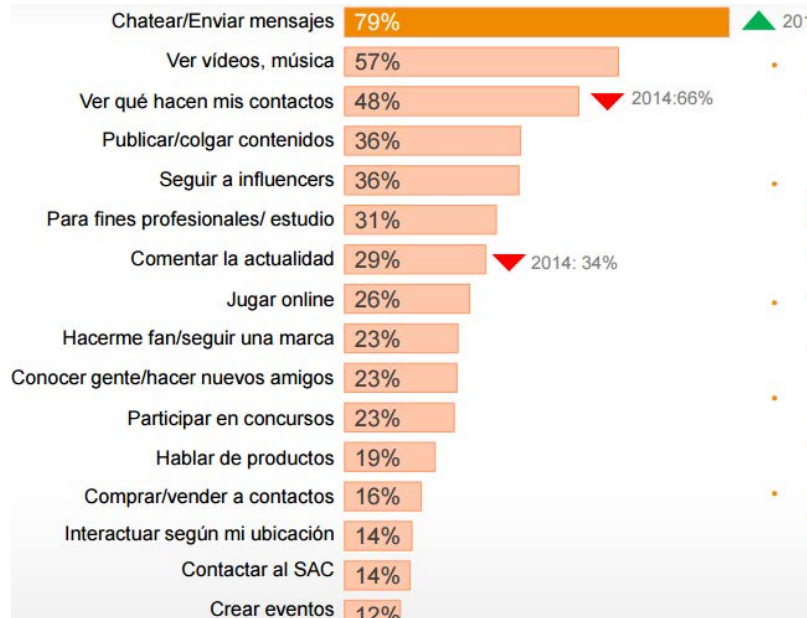


Ilustración Fuente 78

The food industry is in the fifth position in the most followed industry sectors in social networks, with a 25% of the audience. This transformation comes from the hand of the rapid expansion of digital communication media as already discussed in the previous section. Mobile phones and tablets have opened up a new horizon of connectivity and services to a fresh new market.

The prototype presented analyzes social networks for patterns that consumers follow while they are surfing the Web, while they post and enrich social networks. The research is a step towards the preferences of users who use social networks. The market for online products is growing and expectations is to continue growing over the next periods.

The information contained on food retailers pages can help to know habits and consumption patterns recognizing the behavior in relation to seasonability. All this data can help advertising companies and individuals to gather information on the key terms that users post through their comments on social networks.

The reason of the slow development of the project is because I've studied several subjects while I was writing the project, and during the summer I've started working at a consulting firm (PWC). During this period I've acquired various skills while I was studying and working, these skills helped me to develop the project and to incorporate new knowledge.

At an average of one hour a day of work has been devoted a total of 631 hours, this is 31 hours longer than the estimation made with the university policies that should last a set of 24 credits project.

The creation of the company is not very expensive, since a single founding partner can meet the first year of activity of the company, with an initial outlay of € 20000. This amount is low when compared with companies in other sectors as being a service company has no stock and the largest capital of the company is human capital.

The company earns profits from the first year of his birth which makes it an attractive business because its IRR is 26%.

18.5.2 Evaluation of the work carried out

In this project I learned how important it is to study the product well before starting to do any programming task or accounting calculation. This initial study of the product is important for not spinning around, falling into unnecessary mistakes and have delays in product development.

The development of this project has motivated me to improve my knowledge of learning about the different development processes. These development processes throughout the software life I have learned throughout the bachelor study and also in my job.

I have also acquired knowledge of new technologies such as XAMPP and WEKA. These have helped me a lot during development and to check my test results. Also reinforce the knowledge acquired during my bachelor study of the subjects of web programming to develop the system that is part of this project.

I learned a lot of the food retail sector to take advantage in the small business that I've started with a few colleges, for which the development of this project has been good to me dig deeper into this segment of the world of e-commerce sector.

I have found it difficult to scale the entire system, think about its functionalities and choose all the tools. Since I realized that important features are easy to detect yet while others that are complementary are more difficult to detect. However the use of agile methodologies has helped me avoid delays in the system development.

18.5.3 Future work

This section describes some points that can be developed in the future regarding the work which I have already done and raised.

The possibility that the user can download reports as PDF files so he can consult these reports later and more convenient for them to use the tool. A section where customers can advise Facebook pages that can be analysed in the future would be also included.

It also raises that the algorithm can analyse the likes of pages and comments analysed to take into account the popularity of the term being analysed. This idea comes from the possibility that the whole SEO industry and online marketing may happened. This possibility is that the key terms are going to lose some of its strength against the content of the pages.

This possibility may take place because of the appearance of Big data, that has removed the foundations of the positioning of the web pages to be able to check content that is not text

based. For this task, non-relational data bases that are able to store all types of digital content are used.

The digital content of the non-relational data bases can be quickly analyzed by powerful algorithms that provide relevant information to users who are not only saved in text form.

This is why social networks will play a very important role since they offer and also create a lot of content. This content can be valuable form businesses and worth analysing.

19. Anexos

19.1 Bibliografía:

- [1] Wikipedia, PAGERANK, consultada 25/08/2016 Disponible en:
<https://es.wikipedia.org/wiki/PageRank>
- [2] Wikipedia, Posicionamiento en buscadores, consultada el día 19/4/16 Disponible en:
https://es.wikipedia.org/wiki/Posicionamiento_en_buscadores
- [3] Como saber el valor mozRank y la autoridad de una página, un blog o sitio web, norfiPC, consultada 25/8/16 Disponible en: <https://norfipc.com/web/como-saber-conocer-valor-mozrank-pagina-web.php>
- [4] Palabras Clave (Keywords) y Frases Clave (Keyword Phrases), Ninjaseo, consultada 25/8/16 Disponible en: <http://ninjaseo.es/palabras-clave-keywords-frases-clave-keyword-phrase-intro/>
- [5] Search optimization metrics, Wikipedia, consultada 25/8/16 Disponible en:
https://en.wikipedia.org/wiki/Search_engine_optimization_metrics#Ahrefs_URL_Rating_.28AUR.29
- [6] Software, Wikipedia, consultada 26/8/16 Disponible en:
<https://es.wikipedia.org/wiki/Software>
- [7] XAMPP, Wikipedia, consultada 26/8/16 Disponible en:
<https://es.wikipedia.org/wiki/XAMPP>
- [8] Qué es Scrum, Proyectos ágiles, consultada 26/8/16 Disponible en:
<https://proyectosagiles.org/que-es-scrum/>
- [9] Metodologías ágiles para el desarrollo de software: eXtreme Programming (XP), M^o Carmen Penadés y Patricio Letelier consultada 26/8/16 Disponible en:
<http://www.cyta.com.ar/ta0502/v5n2a1.htm>
- [10] Awards club ecommerce, clubecommerce, consultada 25/4/16 Disponible en:
<http://awards.clubecommerce.com/>
- [11] Tiendas alimentación online, tiendas centro red, consultada 26/4/16 Disponible en:
<http://tiendas.centrored.com/online/alimentacion#listadoTiendasBasicas>
- [12] Token, Wikipedia, consultada 26/8/16 Disponible en:
[https://es.wikipedia.org/wiki/Token_\(inform%C3%A1tica\)](https://es.wikipedia.org/wiki/Token_(inform%C3%A1tica))
- [13] Apache Friends Forums, Apache Friends consultada 26/8/16 Disponible en:
<https://www.apachefriends.org/community.html>
- [14] Hacking tutorial: Búsquedas con Google Dorks, Esaú A, consultada 26/8/16 Disponible en: <https://openwebinars.net/hacking-tutorial-busquedas-con-google-dorks/>
- [15] Google Hacking (46 ejemplos), Antonio González, consultada 26/8/16 Disponible en:
<http://antoniogonzalezm.es/google-hacking-46-ejemplos-hacker-contrasenas-usando-google-enemigo-peor/>

- [16] Ventas del sector de la alimentación, Nielsen, consultada 26/8/16 Disponible en: <http://www.nielsen.com/es/es/insights/news/2014/el-ecommerce-de-alimentacion-supone-el-0-6-del-total-de-ventas-del-mercado-de-gran-consumo.html>
- [17] LOS HOGARES QUE COMPRAN ALIMENTACIÓN ONLINE EN ESPAÑA CRECEN EL 25%, Nielsen, consultada 26/8/16 → <http://www.nielsen.com/es/es/press-room/2015/Los-hogares-que-compran-alimentacion-online-en-Espania-crecen-el-25-porciento.html>
- [18] El negocio del súper 'on line' está verde, El país (Fernando Barciela), consultada 26/8/16, Disponible en: http://economia.elpais.com/economia/2015/02/17/actualidad/1424199345_419872.html
- [19] MÁS DE LA MITAD DE LOS CONSUMIDORES YA COMPRA ONLINE VIAJES, ENTRADAS Y ROPA, Nielsen, consultada 26/8/16, Disponible en: <http://www.nielsen.com/es/es/press-room/2015/Mas-de-la-mitad-de-los-consumidores-ya-compra-online-viajes-entradas-y-ropa.html>
- [20] Requisito no funcional, Wikipedia consultada 26/8/16 Disponible en: https://es.wikipedia.org/wiki/Requisito_no_funcional
- [21] Requisito funcional, Wikipedia consultada 26/8/16 Disponible en: https://es.wikipedia.org/wiki/Requisito_funcional
- [22] Proceso de un plan de marketing, imeil, consultada 27/8/16 Disponible en: <http://www.imeil.com.mx/sim/proce/interna.asp>
- [23] Cuanto cobrar por un proyecto sistema informático, Omar Berroterán Silva, consultada 27/8/16 Disponible en: <http://notasprogramacionlkf.blogspot.com.es/2012/01/cuanto-cobrar-por-un-proyecto.html>
- [24] Elegir entre autónomo o sociedad limitada, pymesyautonomos, consultada 27/8/16 Disponible en: <http://www.pymesyautonomos.com/estructura-societaria/elegir-entre-autonomo-o-sociedad-limitada-iii-aspectos-fiscales>
- [25] Capacidad Empresarial y Finanzas, Ingryns, consultada 27/8/16 Disponible en: <http://html.rincondelvago.com/capacidad-empresarial-y-finanzas.html>
- [26] Análisis PEST, Wikipedia, consultada 27/8/16 Disponible en: https://es.wikipedia.org/wiki/An%C3%A1lisis_pest
- [27] Régimen general de la seguridad social, Seguridad social, consultada 28/8/16 Disponible en: http://www.seg-social.es/Internet_1/Trabajadores/CotizacionRecaudaci10777/Regimenes/RegimenGeneraldelas10957/InformacionGeneral/index.htm
- [28] Hosting WordPress 100%, Webempresa, consultada 28/8/16 Disponible en: <https://www.webempresa.com/wordpress/hosting-wordpress.html>
- [29] Incertidumbre e inestabilidad, Silvia Bravo, consultada 28/8/16 Disponible en: http://politica.elpais.com/politica/2016/06/05/actualidad/1465151921_994566.html
- [30] http://contenidos.ceoe.es/CEOE/var/pool/pdf/cms_content_documents-file-348-economia-espanola-conclusiones-positivas-tras-la-supervision-de-la-comision-europea-y-el-bce.pdf, CEOE, consultada 28/6/16 Disponible en:

http://contenidos.ceoe.es/CEOE/var/pool/pdf/cms_content_documents-file-348-economia-espanola-conclusiones-positivas-tras-la-supervision-de-la-comision-europea-y-el-bce.pdf

[31] European Economic Forecast, European comission, consultada 28/8/16 Disponible en: http://ec.europa.eu/economy_finance/publications/eeip/pdf/ip025_en.pdf

[32] El PIB creció el 3,2% en 2015, la mayor subida desde el inicio de la crisis, Alejandro Bolaños, consultada 28/6/16 Disponible en: http://economia.elpais.com/economia/2016/01/29/actualidad/1454053782_817171.html

[33] Producto Interior bruto regional año 2015, consultada 28/8/16 Disponible en: <http://www.ine.es/prensa/np964.pdf>

[34] Desempleo en España, Wikipedia, consultada 28/8/16 Disponible en: https://es.wikipedia.org/wiki/Desempleo_en_Espa%C3%B1a

[35] Previsión del paro en España para 2016 y 2017, Bankinter, consultada 28/8/16 Disponible en: <https://blog.bankinter.com/economia/-/noticia/2015/12/29/prevision-paro-espana-2016-2017.aspx>

[36] Índice de facilidad para hacer negocios, Banco mundial, consultada 28/8/16 Disponible en: http://datos.bancomundial.org/indicador/IC.BUS.EASE.XQ?name_desc=false

[37] La población desciende en España en 2015, Funeraria de guardia, consultada 28/8/16 Disponible en: http://www.funerariadeguardia.com/noticia/95_la-poblacion-desciende-en-espana-en-2015.htm

[38] Life expectancy at birth, OECD, consultada 28/8/16 Disponible en: <https://data.oecd.org/healthstat/life-expectancy-at-birth.htm>

[39] El smartphone se impone por primera vez en España al PC como punto de acceso a la red, marketingdirecto, consultada 28/8/16 Disponible en: <http://www.marketingdirecto.com/digital-general/digital/smartphone-se-impone-primera-vez-espana-al-pc-punto-acceso-la-red/>

[40] Análisis del informe de 2016 del IAB sobre el uso de redes sociales, IAB, consultada 28/8/16 Disponible en: <http://www.antevenio.com/blog/2016/05/analisis-del-informe-de-2016-del-iab-sobre-el-uso-de-redes-sociales/>

[41] El comercio electrónico, a por otro año récord en España, M. Prieto, consultada 28/8/16 Disponible en: <http://www.expansion.com/economia-digital/innovacion/2015/10/25/5626aab5ca4741a9668b467e.html>

[42] El 54% de las pymes ya tiene presencia en las redes sociales, Manuel Moreno, consultada 28/8/16 Disponible en: <http://www.trecebits.com/2015/11/25/el-54-de-las-pymes-ya-tiene-presencia-en-las-redes-sociales/>

[43] Cinco fuerzas de Porter, Wikipedia consultada 28/8/16 Disponible en: https://es.wikipedia.org/wiki/An%C3%A1lisis_Porter_de_las_cinco_fuerzas#.28F4.29_Amenaza_de_productos_sustitutos

[44] Usuarios activos en redes sociales en 2016, SoMaMFyC, consultada 29/8/16 Disponible en: <https://nuevastecsomamfyc.wordpress.com/2016/02/26/usuarios-activos-en-redes-sociales-en-2016/>

[45] Historia, evolución y muerte del Pagerank, PasionSEO, consultada 29/8/16 Disponible en: <http://www.pasionseo.com/historia-evolucion-y-muerte-del-pagerank/>

[46] Posicionamiento en buscadores, Wikipedia, consultada 29/8/16 Disponible en: https://es.wikipedia.org/wiki/Posicionamiento_en_buscadores

[47] La importancia del análisis de palabras clave, Adrián, consultada 29/8/16 Disponible en: <http://www.lapocima.es/posicionamiento-web-analisis-de-palabras-clave/>

[48] Rendimiento de Windows 10 frente a Windows 7, Juan Ranchal, consultada 29/8/16 Disponible en: <http://www.muycomputer.com/2015/08/20/rendimiento-de-windows-10>

[49] Análisis Porter de las cinco fuerzas, Wikipedia, consultada 29/8/16 Disponible en: https://es.wikipedia.org/wiki/An%C3%A1lisis_Porter_de_las_cinco_fuerzas

[50] Las Mejores Herramientas SEO para 2016, Javier Marcilla, consultada 30/8/16 Disponible en: <http://ninjaseo.es/herramientas-seo-para-posicionar-en-google/>

[51] Herramientas SEO para posicionar en Google, Javier Marcilla, consultada 30/8/16 Disponible en: <http://ninjaseo.es/herramientas-seo-para-posicionar-en-google/#Bufferapp>

[52] Cómo se hace un análisis DAFO, emprendedores, consultada 30/8/16 Disponible en: <http://www.emprendedores.es/gestion/como-hacer-un-dafo/como-hacer-un-dafo2>

[53] Objetivos estratégicos, Wikipedia, consultada 31/8/16 Disponible en: https://es.wikipedia.org/wiki/Objetivos_estrat%C3%A9gicos

[54] Definición de escudo fiscal, Gonzalo Barroso, consultada 31/8/16 Disponible en: <http://empresayeconomia.republica.com/contabilidad/definicion-de-escudo-fiscal.html>

[55] El plan de marketing, Educamarketing, consultada 31/8/16 Disponible en: <http://educamarketing.unex.es/Docs/guias/gu%C3%ADa%20para%20la%20elaboraci%C3%B3n%20de%20un%20plan%20de%20marketing.pdf>

[56] Plan de marketing, Roberto Espinosa, consultada 31/8/16 Disponible en: <http://robertoespinosa.es/2014/03/25/como-elaborar-el-plan-de-marketing/>

[57] Producto (marketing), Wikipedia, consultada 31/8/16 Disponible en: [https://es.wikipedia.org/wiki/Producto_\(marketing\)#Clasificaci.C3.B3n_de_los_productos_de_consumo_y_servicio](https://es.wikipedia.org/wiki/Producto_(marketing)#Clasificaci.C3.B3n_de_los_productos_de_consumo_y_servicio)

[58] Etapas en la vida de un producto, Debitoor, consultada 31/8/16 Disponible en: <https://debitoor.es/glosario/definicion-cvp>

[59] Internet es el canal más rentable en comunicación pero todavía tiene mucho que hacer en la generación de ventas, Marketing directo, consultada 31/8/16 Disponible en: <http://www.marketingdirecto.com/digital-general/digital/internet-es-el-canal-mas-rentable-en-comunicacion-pero-todavia-tiene-mucho-que-hacer-en-la-generacion-de-ventas/>

[60] Internet como canal de ventas, Okhosting, consultada 31/8/16 Disponible en: <http://okhosting.com/blog/el-internet-como-un-nuevo-canal-de-comunicacion/>

[61] Plan de recursos humanos, ceei, consultada 1/9/16 Disponible en: <http://www.guia.ceei.es/interior.asp?MP=8&MS=9>

- [62], Organigrama, Wikipedia, consultada 1/9/16 Disponible en:
<https://es.wikipedia.org/wiki/Organigrama>
- [63] Los 3 sistemas de retribución más eficaces, Retos directivos, consultada 1/9/16
Disponible en: <http://retos-directivos.eae.es/los-tres-sistemas-de-retribucion-mas-eficaces/>
- [64] Las retribuciones extrasalariales, Susana Palacios, consultada 1/9/16 Disponible en:
http://cincodias.com/cincodias/2001/08/06/empresas/997190084_850215.html
- [65] Plan económico y financiero, cámara Santa Cruz de Tenerife, consultada 1/9/16
Disponible en: <http://www.creacionempresas.com/plan-de-viabilidad/que-es-un-plan-de-empresa-viabilidad/plan-economico-y-financiero>.
- [66] Arquitectura de las aplicaciones web, Instituto tecnológico de Matehuala, consultada 3/9/16 Disponible en: <https://programacionwebisc.wordpress.com/2-1-arquitectura-de-las-aplicaciones-web/>
- [67] Clave primaria, Wikipedia, consultada 3/9/16 Disponible en:
https://es.wikipedia.org/wiki/Clave_primaria
- [68] Pruebas aceptación del cliente, Yanet Fernández, consultada 3/9/16 Disponible en:
<http://www.monografias.com/trabajos36/pruebas-de-aceptacion/pruebas-de-aceptacion2.shtml>
- [69] Localizador de recursos uniforme, Wikipedia, consultada 4/9/16 Disponible en:
https://es.wikipedia.org/wiki/Localizador_de_recursos_uniforme
- [70] Futuro del marketing, marketingonline, consultada 7/9/16 Disponible en:
<http://www.marketingdirecto.com/marketing-general/publicidad/la-inversion-publicitaria-crecera-46-2016-segun-zenithoptimedia/>
- [71] Inversión Publicitaria en Internet, ymedia, consultada 7/9/16 Disponible en:
<http://www.ymedia.es/es/canal-ymedia/otros-medios/inversion/215/24-feb-2016/estudio-infoadex-de-la-inversion-publicitaria-en-espana-2015>
- [72] Evolución PIB español, El país, consultada 7/9/16 Disponible en:
http://ep01.epimg.net/economia/imagenes/2015/12/22/actualidad/1450776874_178140_1450799579_sumario_normal.png
- [73] Estudio Inversión Publicitaria, IAB, consultada 7/9/16 Disponible en:
http://www.iabspain.net/wp-content/uploads/downloads/2016/03/Estudio_Inversi%C3%B3n_Publicitaria_Medios_Digitales_2015_IAB_Spain_vReducida.pdf
- [74] Bono español a 5 años, El economista, consultada 11/09/16 Disponible en:
<http://www.eleconomista.es/bono/bono-espana-5-years>
- [75]¿Qué es el ROI?, 40 de fiebre, consultada 11/9/16 Disponible en:
<https://www.40defiebre.com/que-es/roi/>
- [76] Rentabilidad financiera, Wikipedia, consultada 11/9/16 Disponible en:
https://es.wikipedia.org/wiki/Rentabilidad_financiera

[77] Uso de redes sociales, El tipómetro, consultada 13/9/16 Disponible en:
<http://franbarquilla.com/estudio-sobre-el-estado-de-internet-y-las-redes-sociales-en-2016/>

[78] Estudio redes sociales, IAB, consultada 13/9/16 Disponible en:
http://www.iabspain.net/wp-content/uploads/downloads/2016/04/IAB_EstudioRedesSociales_2016_VCorta.pdf

[79] API de Facebook, Facebook, consultada 14/9/16 Disponible en:
<https://developers.facebook.com/tools/debug/accesstoken/?q=EAABnZBHI1WMQBAE2QnFPJin5RgMWT3ZB8WmPsKeMwquM16FOd243VEvuafFN1esYK2Ft4EtnAsTeZCLNWSzgFQWEbw9JT uMHqtOION9I99WQS6vHFYIOLLAWATfu55UTSXmkPARRszGB6654kNpjh59prTc435ZAZAyqjPrzZARwZDZD>

[80] Normas de Facebook, Facebook consultada 14/9/16 Disponible en:
<https://developers.facebook.com/policy/>

[81] Ley orgánica de protección de datos, Agencia española de protección de datos, Disponible en: <https://www.agpd.es/portaIwebAGPD/canaIdocumentacion/legislacion/estatal/index-ides-idphp.php>

19.2 Glosario de términos y acrónimos.

Autoridad de dominio: empleado por SEOmoz para expresar numéricamente la importancia y el prestigio de un dominio y de una página específica. El valor depende de varios factores como el tiempo de funcionamiento, la presencia o ausencia de spam y malware, la autoridad del servicio de hosting, el tiempo de existencia del dominio, el periodo de pago adelantado del dominio y muchos otros. [3]

Taxonomía: la taxonomía se puede definir como una ciencia para el estudio de la clasificación. En este caso se ha utilizado para clasificar y ordenar en categorías los distintos términos clave del sector de la alimentación.

MOZRank: método alternativo creado para catalogar la importancia y popularidad de las páginas web de internet. Sigue el mismo objetivo del PageRank de Google, pero la gran diferencia consiste en que el PageRank se aplica solo a las páginas que se encuentran en el índice de Google, mientras que el mozRank (mR) y el mozTrust (mT), término algo similar, se aplican a cualquier página de internet, o sea son de carácter global.[3]

Software: *al equipo lógico o soporte lógico* de un sistema informático, que comprende el conjunto de los componentes lógicos necesarios que hacen posible la realización de tareas específicas, en contraposición a los componentes físicos que son llamados hardware.[6]

XAMPP: es un servidor independiente de plataforma, software libre, que consiste principalmente en el sistema de gestión de bases de datos MySQL, el servidor web Apache y los intérpretes para lenguajes de script: PHP y Perl. El nombre proviene del acrónimo de X (para cualquiera de los diferentes sistemas operativos), Apache, MySQL, PHP, Perl. Desde la versión "5.6.15", XAMPP cambió la base de datos de MySQL A MariaDB. El cual es un fork de MySQL con licencia GPL.[7]

Token: también llamado componente léxico es una cadena de caracteres que tiene un significado coherente en cierto lenguaje de programación. [12]

SEO: (search engine optimization) es el proceso técnico mediante el cual se realizan cambios en la estructura e información de una página web, con el objetivo de mejorar la visibilidad de un sitio web en los resultados orgánicos de los diferentes buscadores. [45]

Escudo fiscal: Según el Instituto de Estudios Superiores de la Empresa (IESE), “el escudo fiscal es la estrategia para reducir los impuestos mediante desgravaciones fiscales” [54].

PWC: PricewaterhouseCoopers

URL: Uniform Resource Identifier o localizador de recursos uniforme, sirve para localizar direcciones de recursos en la web. Estos recursos a los que apunta pueden ser variables en el tiempo [69].

19.3 Tablas anexas de cuentas

P&L	2016	2017	2018
Ingresos	23000	24311	25939,84
Aprovisionamientos	-1185	-1185	-1185
Gastos Personal	-19350	-19350	-19350
Personal	-15000	-15000	-15000
Cargas Sociales	-4350	-4350	-4350
Amortizaciones	-166	-166	-166
Beneficio Operativo	2299	3610	5238,837
Impuesto Sociedades (30%)	-689,7	-1083	-1571,65
Resultado Neto	1609,3	2527	3667,186

Tabla 106. Cuenta de pérdidas y ganancias 16 al 18

	2016	2017	2018
Ingreso:	43000	24311	25939,84
Capital	20000		
Ventas	23000	24311	25939,84
Salidas:	-21531	-20205,7	-20599
Aprovisionamientos	-1185	0	0
Gastos Personal	-19350	-19350	-19350
Inversiones	-830	0	0
Impuesto de sociedades	0	-689,7	-1083
Amortizaciones	-166	-166	-166

Cash flow	21469	4105,3	5340,837
-----------	-------	--------	----------

Tabla 107. Entradas y salidas de efectivo, 16 al 18

Inversiones		5 años de vida útil
Móvil	100	20
Ordenador	500	100
Silla	50	10
Mesa	30	6
Impresora	50	10
Licencias	100	20
Total	830 €	166 €
Personal		
Sueldo bruto	15000	
Cotizaciones a cargo de la empresa	4350	Lo saco de la página [27] La cotización se ha sacado social
Gtos de personal	19350	
Otros gastos		
Cuenta Dropbox	600	
Papel	6	
Luz	200	
Agua	200	
Tinta impresora	30	
Hosting	99	
Varios	50	
Total	1185	Todos los años tengo es
Diferencia sin el coste de DB	585	

Tabla 108. Datos de los gastos de la empresa

		2016	2017	2018
Activo no corriente	ANC	664	498	332

Inmovilizado intangible	Inmovilizado Intangible	100	100	100
Amortización acumulada Inmovilizado intangible	A.A Inmov Intan	-20	-40	-60
Inmovilizado material	Inmovilizado material	730	730	730
Amortización acumulada Inmovilizado material	A.A Inmov material	-146	-292	-438
Activo corriente	AC	21635	21635	29043,14
Tesorería	Tesorería	21635	24721,3	29043,14
Total		22299	25219,3	29375,14

		2016	2017	2018
Pasivo neto	PN	21609,3	24136,3	27803,49
Capital	Capital	20000	21609,3	24136,3
Resultado	Resultado	1609,3	2527	3667,186
Reserva legal		0	0	160,93
Reserva voluntaria		0	0	1448,37
Pasivo no corriente	PNC	0	0	0
Pasivo corriente	PC	689,7	1083	1571,651
Pasivos con hacienda	Pasivos con Montoro	689,7	1083	1571,651
Total		22299	25219,3	29375,14

Tabla 109. Balance de cuentas

Beneficios	I. inicial	2016	2017	2018	n años	k (tasa de interés)
	-20000	1609,3	2527	3667,186	3	0,085 %
		k=	0,00085			
		VAN=	6500,869			

Tabla 110. Tabla de justificación del VAN

Beneficios para los próximos 10 años.	versión inicial	2016	2017	2018	2019	2020	2021	2022	2023
	-20000	23000	24311	25939,84	27677,81	29532,22	31510,88	33622,11	35874,79
os		-1185	-1185	-1185	-1185	-1185	-1185	-1185	-1185
		-19350	-19350	-19350	-19350	-19350	-19350	-19350	-19350
		-15000	-15000	-15000	-15000	-15000	-15000	-15000	-15000
		-4350	-4350	-4350	-4350	-4350	-4350	-4350	-4350
		-166	-166	-166	-166	-166	-166	-166	-166
vo		2299	4210	5838,837	6976,806	8831,219	10809,88	12921,11	15173,79
des (30%)		-689,7	-1263	-1751,65	-2093,04	-2649,37	-3242,96	-3876,33	-4552,14
	-20000	1609,3	2947	4087,186	4883,764	6181,853	7566,914	9044,775	10621,65

TIR 0,244572

TIR a los 10 años.

Tabla 111. Tabla de justificación de la TIR

Balance ene-16	Año 0	
ANC	830	PN 20000
Inmovilizado Intangible	100	Capital 20000
Inmovilizado material	730 €	PNC 0
AC	19170	
Tesorería	19170	PC 0
Total	20000	20000

Tabla 112. Tabla de balance de ene-16

EFE	Estado de flujos de caja
Capital	20000
inversiones	-830
Tesorería	19170

Tabla 113. Justificación de estados de flujo de caja ene-16

				El precio de 8 € está basado en la competencia, ya que esta se mu			
Términos incluidos	Uds	1150		Precio	8	Ingreso Estimado	9200
Nuevos ámbitos	Uds	6		Precio	2300	Ingreso Estimado	13800
Total		1156					23000

Justificación de los ingresos:

Para añadir nuevos ámbitos a la aplicación, como puede ser, una sección dentro de la alimentación que no exista u otro campo o clase genérica distinto. El sector del SEO se organiza con el estándar que siguen las empresas de consultoría. Las consultoras cobran entre 40 y 60€ la hora dependiendo del proyecto. En el caso de LULAF el precio por hora es de 50€.

Nº horas 46 Horas que necesito para añadir cada ámbito nuevo

Estas horas son las que ha estimado el desarrollador para añadir cada ámbito por su experiencia desarrollando el sistema. Entonces $46 \times 50 = 2300€$

Tabla 114. Justificación de los ingresos en 16

Fórmula utilizada							
Q= CF/ (Pº - CV)	Pto muerto en conj	2710,89425	aprox	2710,89425	Unidades		Pº= precio
Desglose del punto muerto por cada uno de los productos que se ofrecen. Tiene más sentido verlo así.							
	Términos incluidos	2725,0875	aproximado	2726	Unidades		CF= costes fijos
	Nuevos ámbitos	9,47856522	aproximado	10	Unidades		CV= costes variab
(Coste Fijo total* Proporción de coste de TÉRMINOS INCLUIDOS/ precio TÉRMINOS INCLUIDOS) + (Coste Fijo total* Proporción de coste de NUEVC							
proporción de unidades vendidas en:							
	Términos incluidos:		0,99480969				
	Nuevos ámbitos		0,00519031				
	Comprobación		1		Tiene que dar 1.		

Tabla 115. Cálculo del punto muerto justificado.

ROI= 20,77312614

ROE= 13,18966231

Tabla 116. Justificación del ROE y del ROI.

Balance dic-16	Año 0	
ANC	664	PN 21609,3
Inmovilizado Intangible	100	Capital 20000
A.A Inmov Intan	-20	Resultado 1609,3
Inmovilizado material	730	PNC 0
A.A Inmov material	-146	
AC	21635	
Tesorería	21635	PC 689,7
		Pasivos con hacienda 689,7
Total	22299	22299

Tabla 117. Balance en 2016

EFE	
Resultado	1609,3
Flujos de actividades Operativas	689,7
Pasivos con hacienda	689,7
Flujos de Financiación	20000
Creación de la empresa	20000
Otros	
Flujos	-664
Amortizaciones	166

inversiones	-830
Total efectivo	21635

Tabla 118. Estado de flujos de efectivo dic-16

	Balance dic-17	Año 0		
	ANC	498	PN	24136,3
	Inmovilizado			
	Intangible	100	Capital	21609,3
Amortización acumulada	A.A Inmov Intan	-40	Resultado	2527
	Inmovilizado material	730 €	Reserva legal	160,93
	A.A Inmov material	-292	Reserva voluntaria	1448,37
		24.721		
	AC	€	PNC	0
		24.721		
	Tesorería	€	PC	1083
			Pasivos con Hacienda	1083
Total		25219,3		25219,3
				0

Tabla 119. Balance de dic -17

EFE	
Resultado	2527
Flujos de actividades Operativas	393,3
Pasivos con Hacienda	1083
Pagos a Hacienda	-689,7
Flujos de Financiación	0
Flujos de Ingreso	
Otros	
Flujos	166
Amortizaciones	166
Total efectivo	3086,3
-21635	
Tesorería inicial	21635
Variaciones	3086,3
Tesorería final	24721,3

Tabla 120. Estado de flujos de efectivo de di-17

	Balance dic-18	Año 0	
	ANC	332	PN 27803,4859
	Inmovilizado Intangible	100	Capital 24136,3
Amortización acumulada	A.A Inmov Intan	-60	Resultado 3667,1859
	Inmovilizado material	730	Reserva legal 413,63
	A.A Inmov material	-438	Reserva voluntaria 3722,67
	AC	29043,137	PNC 0
	Tesorería	29.043 €	PC 1571,6511
			Pasivos con Montoro 1571,6511
Total		29375,137	29375,137

Tabla 121. Balance de dic-18

	EFE	
	Resultado	3667,1859
Flujos de actividades Operativas		488,6511
	Pasivos con hacienda	1571,6511
	Pagos a hacienda	-1083
Flujos de Financiación		0
Otros Flujos		166
	Amortizaciones	166 €
Total variaciones		4321,837
Tesorería inicial (2017)	24.721 €	
Variaciones de Tesorería	4321,837	
Tesorería final	29.043 €	

Tabla 122. Estado de flujos de efectivo dic-18