

Seobilidad. Cómo aplicar el SEO centrado en la Usabilidad.

Memoria de Proyecto Final de Grado

Grado Multimedia

Usabilidad e Interfaces

Autor: Patricia Luengo Carretero

Consultor: Judit Casacuberta Bagó

Profesor: Ferrán Giménez Prado

10/01/2022

Créditos/Copyright



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-SinObraDerivada [3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

Abstract

En este TFG tratamos de contestar a las siguientes preguntas ¿deberíamos diseñar pensando en el usuario o pensando en el buscador?, ¿cuánto influye la usabilidad de un sitio web en su posicionamiento en buscadores?, ¿podemos cuantificarlo?, ¿es posible aplicar el SEO y la usabilidad en conjunto?

Por un lado, la revisión de los contenidos teóricos de la usabilidad y la recopilación de acciones que se llevan a cabo en las estrategias de posicionamiento han revelado que tienen muchos aspectos en común sobre todo en las etapas de análisis, modelado del usuario, arquitectura de la información y diseño de contenidos.

Por otro lado, la investigación práctica basada en la recopilación de numerosos indicadores y la implementación de un modelo de análisis matemático de regresión múltiple han revelado que hay muchos factores relacionados con la usabilidad que tienen influencia en el posicionamiento pero que aún están lejos de tener la relevancia de los factores exclusivamente relacionados con el SEO, es decir, los enlaces entrantes o *backlinks*.

Finalmente, se llega a la conclusión de que la usabilidad y el SEO deben dejar de utilizarse por separado a la hora de diseñar un sitio web ya que juntas se enriquecen y nos ayudan a tener una visión más amplia de quienes son los usuarios.

Palabras clave: SEO, usabilidad, seobilidad, DCU, posicionamiento, diseño web, arquitectura de la información, modelado de usuarios, diseño de contenido, búsqueda, enlaces entrantes

Abstract (english version)

In this TFG we try to answer the following questions: Should we design thinking about the user or thinking about the search engine? How much does the usability of a website influence its search engine positioning? Can we quantify it? Is it possible to apply the SEO and usability together?

On the one hand, the review of the theoretical contents of usability and the actions that are carried out in SEO strategies have revealed that they have many aspects in common, especially analysis, user modeling, architecture of information and content design.

On the other hand, practical research based on numerous indicators and the implementation of a multiple regression mathematical analysis model have revealed that there are many usability factors that influence positioning but are still far from having the same relevance of those factors exclusively related to SEO, that is, backlinks.

Finally, it is concluded that usability and SEO should stop being used separately when designing a website since the boards are enriched and help us to have a broader vision of who the users are.

Keywords: SEO, usability, seobility, DCU, positioning, web design, information architecture, user modeling, content design, search, backlinks

Índice

1. Motivación.....	8
2. Descripción	9
3. Objetivos.....	10
3.1 Principales.....	10
3.2 Secundarios	10
4. Contenidos.....	11
5. Planificación.....	12
6. Marco teórico	13
6.1 Usabilidad	13
6.2 SEO.....	28
6.3 Combinación Usabilidad y SEO	42
6.4 Herramientas de usabilidad.....	42
6.5 Herramientas SEO	43
7. Marco práctico	46
7.1 Datos de partida.....	46
7.2 Selección del sector	46
7.3 Estudio de competidores	47
7.4 Análisis de palabras clave	51
7.5 Métricas de la muestra elegida.....	55
7.6 Modelo empleado y resultados del análisis	66
8. Metodología para Seoabilidad.....	76
8.1. Análisis.....	77
8.2. Modelado del usuario.....	78
8.3. Diseño.....	82
8.4. Prototipo.....	87
8.6. Evaluación.....	87
8.7. Seguimiento	88
8.8. Métricas.....	90
9. Recomendaciones	92
10. Conclusiones	95
Anexo 1. Entregables del proyecto.....	97
Anexo 2. Librerías.....	98
Anexo 4. Bibliografía.....	99

Figuras y tablas

Índice de figuras

Ilustración 1. Planificación.....	12
Ilustración 2. Factores clave en usabilidad.....	14
Ilustración 3. Principios básicos de usabilidad	15
Ilustración 4. Principios generales de usabilidad.....	17
Ilustración 5. Principios heurísticos de diseño de interacción de Tognazzini	20
Ilustración 6. Diseño centrado en el usuario (DCU)	22
Ilustración 7. Fases del proceso de indexación.....	29
Ilustración 8. Intención de búsqueda.....	30
Ilustración 9. <i>Searcher</i> persona.	32
Ilustración 10. Viaje del cliente.....	33
Ilustración 11. Metodología SEO.....	35
Ilustración 12. Vista de ejemplo de competidores del sitio web <i>piensoymascotas.com</i>	48
Ilustración 13. Preparación del entorno en <i>Python</i> y <i>Google Colab</i>	49
Ilustración 14. Carga de datos en <i>Python</i>	49
Ilustración 15. Función que recopila todos los competidores.....	50
Ilustración 16. Muestra parcial de todos los competidores encontrados.....	50
Ilustración 17. Vista de ejemplo de obtención de palabras clave en el dominio <i>bitiba.es</i>	52
Ilustración 18. Carga de datos de palabras clave de la muestra de dominios competidores.....	52
Ilustración 19. Comprobación de la carga de datos.....	53
Ilustración 20. Función que busca las palabras clave en común entre todos los sitios web.....	53
Ilustración 21. Obtención del ranking para la búsqueda "tiendas de perros"	54
Ilustración 22. Ranking de los diez sitios web a auditar	55
Ilustración 23. Datos de estimación de CTR	58
Ilustración 24. Ejemplo de captura de datos de velocidad móvil con <i>Test my site</i>	61
Ilustración 25. Ejemplo de captura de datos de WPO.....	62
Ilustración 26. Vista parcial de la carga de datos en <i>Python</i>	66
Ilustración 27. Matriz de correlación.....	67
Ilustración 28. Variables predictoras.....	69
Ilustración 29. Normalización variables predictoras.....	70
Ilustración 30. Variable respuesta.....	71
Ilustración 31. Implementación del modelo.....	71
Ilustración 32. Obtención de intercept, coeficientes de regresión y valor de R.....	72
Ilustración 33. Datos estadísticos de las variables predictoras.....	73
Ilustración 34. Vista parcial de datos estadísticos.....	73
Ilustración 35. Predicción.....	74
Ilustración 36. Ejemplo de una ficha persona.....	79
Ilustración 37. Ejemplo de <i>journey map</i>	81
Ilustración 38. Seoabilidad.....	89

Índice de tablas

Tabla 1. Listado sectorial Mascotas.	47
Tabla 2. Listado ecommerce sector mascotas.	47
Tabla 3. Porcentaje de url amigables	56
Tabla 4. Porcentaje de páginas con <i>thin content</i> .	57
Tabla 5. Porcentaje de enlaces	57
Tabla 6. Porcentaje de páginas correctamente etiquetadas	58
Tabla 7. Autoridad de dominio.	59
Tabla 8. Páginas por visita	59
Tabla 9. Duración media de la visita	60
Tabla 10. Tasa de rebote	60
Tabla 11. Velocidad móvil	61
Tabla 12. Enlaces totales	64
Tabla 13. Enlaces follow	64
Tabla 14. Enlaces no follow	65
Tabla 15. Dominios únicos de referencia	65

1. Motivación

Desde 2009 tengo una tienda online que se ha ido transformando en función de muchos factores como la evolución de Internet, el uso de distintos *CMS (Content Management System)* como *Joomla*, *Prestashop* y *Wordpress*, mi propio desarrollo personal realizando todo tipo de cursos junto con el Grado Multimedia y como no, los cambios en el algoritmo de *Google*.

Al principio resultó sencillo posicionarse en *Google*, no había demasiada competencia en mi ámbito y con llenar el *snippet* de las páginas de la web con multitud y variopintas combinaciones de palabras clave era más que suficiente.

Con los años, el algoritmo de *Google* se fue refinando y el *SEO* se fue complicando en la misma medida que la competencia crecía sin parar.

A lo largo de todos estos años mi tienda se ha mantenido férrea entre las tres primeras posiciones de *Google* para diversas búsquedas, junto a ella han ido apareciendo competidores que por distintas razones fueron expulsados de la primera página del buscador después de algún *update* de *Google*, algunos con una experiencia de usuario dudosa y muchos enlaces entrantes, otros con buen contenido y uso de muchas palabras clave y otros que han cesado por sí mismos porque la conversión no era la esperada.

Todo ello me ha llevado a preguntarme ¿por qué mi tienda no ha sufrido estos bailes de posiciones en tanto tiempo? La respuesta más sencilla para mí es que siempre he pensado en qué le resultaba más fácil al usuario a la hora de utilizar mi web, al principio sin conocimiento alguno y usando el sentido común y hoy en día implementando todo lo que he aprendido hasta ahora.

Entonces, si creo firmemente en que la usabilidad ha sido un punto clave para mi proyecto, sin obviar el *SEO* y tratando siempre de buscar el equilibrio ¿esta combinación puede ser la clave para el futuro del posicionamiento en el buscador? Esta es mi motivación para abordar este TFG.

2. Descripción

El *SEO* (*Search Engine Optimization*), es decir, la optimización en motores de búsqueda y la usabilidad, son técnicas que se implementan sobre los sitios web de forma independiente.

Con el *SEO* tratamos de aparecer en las primeras posiciones del buscador realizando diferentes acciones sin tener en cuenta si el usuario que haga clic en el resultado tendrá una buena o mala experiencia en el sitio. Con la usabilidad, el objetivo es que al usuario le resulte fácil utilizar la web una vez está navegando en ella.

Lo que se puede afirmar es que una buena implementación del *SEO* por sí sola no asegura que el sitio web sea usable, y al revés, que un sitio tenga una buena usabilidad por sí sola no va a situarnos en las primeras posiciones del buscador, de hecho, puede ocurrir que ni siquiera aparezcamos en él.

La cuestión principal que aquí se plantea es: ¿podemos mejorar nuestro ranking mejorando aquellos aspectos en los que la usabilidad y el *SEO* están estrechamente relacionados? ¿podemos implementar una forma de trabajo que combine ambas disciplinas a pesar de que se apliquen por separado? La hipótesis de la que se parte es que sí, que implementar *SEO* centrado en la usabilidad y al revés, mejorará el posicionamiento.

Existe poca literatura sobre esta cuestión, ya que el estudio de la usabilidad se imparte mayoritariamente en el ámbito académico mientras que el *SEO* no tiene tanta presencia y eso ha hecho que haya pocas investigaciones que los relacionen.

En este TFG se pretende sentar unas bases teóricas y prácticas para que ambas disciplinas dejen de implementarse por separado.

3. Objetivos

El objetivo principal de este TFG es conseguir demostrar como la usabilidad y el *SEO*, tradicionalmente aplicados por separado, pueden mejorar la visibilidad de los sitios web si se hace una implementación conjunta.

3.1 Principales

Los objetivos principales del TFG son:

- Analizar la relación en el posicionamiento aplicando técnicas de *SEO* y usabilidad.
- Identificar los factores clave para el análisis conjunto de ambas técnicas.
- Analizar cuantitativamente la mejora en el posicionamiento gracias al tándem *SEO* y usabilidad.
- Elaborar una metodología de implementación.

3.2 Secundarios

Objetivos adicionales que complementan el TFG.

- Enumerar las principales herramientas a utilizar en el análisis del *SEO* y la usabilidad.
- Describir una serie de recomendaciones.
- Introducción en el análisis de algoritmos de inteligencia artificial.

4. Contenidos

En primer lugar, se realizará un trabajo teórico contextualizando, por un lado, la usabilidad y sus principios y por otro, el *SEO* y sus principales aspectos, para finalmente definir los factores en común que tienen ambas disciplinas para mejorar el posicionamiento de un sitio web. Esta combinación podríamos definirla como seobilidad.

Enumeraremos las principales herramientas utilizadas para aplicar el *SEO* y la usabilidad que nos permitan identificar cuáles y en que parte de toda la metodología han de usarse.

Una vez definidas las bases teóricas y las herramientas, se implementará una parte práctica en la que se analizarán varios sitios web con los factores extraídos de la parte teórica.

Para llevar a cabo este análisis se utilizará una base de datos de sitios web agrupados por sectores y clasificados dentro de cada sector por orden de relevancia en *Google* después del último *Core Update* de Julio de 2021. (Cachón, 2021)

Se elegirá un sector y los diez sitios web más relevantes y se realizará un estudio de palabras clave para extraer la búsqueda que sea más relevante y que todas las webs tengan en común, es decir, trataremos de elegir la palabra más competida. Una vez realizado esto definiremos la posición de cada una en el buscador.

Después pasaremos a analizar cada una de las webs desde las bases teóricas a las que hemos llegado para tratar de encontrar una correlación entre las posiciones de los sitios web y los factores de usabilidad y seo y poder extraer conclusiones. Para este análisis se usará un modelo matemático que ofrezca una base científica al análisis.

Finalmente, definiremos una nueva metodología que combine ambas disciplinas y haremos una serie de recomendaciones a tener en cuenta para implementarla.

5. Planificación

Planificación TFG							
TAREA	Inicio	Fin	Septiembre	Octubre	Noviembre	Diciembre	Enero
PEC1	15/09/21	28/09/21	14 días				
Tarea 0. Idea.	15/09/21	19/09/21	5 días				
Tarea 1. Secciones previas	20/09/21	22/09/21	3 días				
Tarea 2. Introducción	23/09/21	23/09/21	1 día				
Tarea 3. Descripción	24/09/21	24/09/21	1 día				
Tarea 4. Objetivos	25/09/21	25/09/21	1 día				
Tarea 5. Contenidos	26/09/21	26/09/21	1 día				
Tarea 6. Planificación	27/09/21	28/09/21	2 días				
PEC2	29/09/21	27/10/21	29 días				
Tarea 7. Usabilidad	29/09/21	9/10/21	2 días	9 días			
Tarea 8. SEO	10/10/21	20/10/21		11 días			
Tarea 9. Herramientas SEO	21/10/21	24/10/21		4 días			
Tarea 10. Herramientas de usabilidad	25/10/21	27/10/21		3 días			
PEC3	28/10/21	28/11/21		32 días			
Tarea 11. Datos y selección del sector	28/10/21	30/10/21		3 días			
Tarea 12. Análisis de competidores	30/10/21	31/10/21		2 días			
Tarea 13. Análisis de palabras clave	1/11/21	2/11/21			2 días		
Tarea 14. Análisis Seobilidad	3/11/21	14/09/21			12 días		
Tarea 15. Modelo	15/09/21	26/11/21			12 días		
Tarea 16. Resultados	27/11/21	28/11/21			1 día		
PEC4	1/12/21	10/01/22				42 días	
Tarea 17. Metodología Seobilidad	1/12/21	13/12/21				13 días	
Tarea 18. Errores	14/12/21	16/12/21				3 días	
Tarea 19. Conclusiones	16/12/21	19/12/21				4 días	
Tarea 20. Finalización de la memoria	20/12/21	27/12/21				8 días	
Tarea 21. Presentación escrita	28/12/21	2/01/22				4 días	2 días
Tarea 22. Elaboración del vídeo	3/1/22	8/1/22					6 días
Tarea 23. Autoinforme	9/1/22	9/1/22					1 día
Tarea 24. Publicación del TFG	10/01/22	10/01/22					1 día

Ilustración 1. Planificación

6. Marco teórico

6.1 Usabilidad

6.1.1. Definición

La usabilidad es un término muy utilizado en el ámbito de las nuevas tecnologías que sin embargo no recoge la RAE y que muchos autores han tratado de definir desde distintos puntos de vista; se ha escogido la que aporta Yusef Hassan Montero en su libro *Experiencia de Usuario: Principios y Métodos*, por su sencillez y sentido global.

“La usabilidad es un atributo de calidad de un producto que se refiere sencillamente a su facilidad de uso. No se trata de un atributo universal, ya que un producto será usable si lo es para su audiencia específica y para el propósito específico con el que fue diseñado.” (Montero, 2015).

6.1.2. Factores clave

La usabilidad tiene una serie de factores clave o dimensiones que fueron descritos por Jakob Nielsen (Nielsen, *10 Usability Heuristics for User Interface Design*, 1995) y que pueden desglosarse en criterios de medición objetivos y subjetivos:

- **Criterios objetivos:** Aquellos donde evaluamos al usuario a través de la observación mientras utiliza el producto, entre los que tenemos:
 - **Aprendizaje.** Facilidad con la que el usuario aprende a utilizar el producto, cuanto más fácil le resulte más rápido podrá empezar a usarlo.
 - **Eficiencia.** Tiempo que emplea el usuario en realizar cualquier acción para la que ha sido diseñado el producto una vez que ha aprendido a usarlo.
 - **Memorización.** Se hace una distinción entre los usuarios nuevos y recurrentes, éstos deberán utilizar el producto con más facilidad y rapidez.

- **Eficacia.** El producto deberá estar pensado para que no genere errores y en el caso de que lo haga éstos sean mínimos además de ofrecer al usuario una salida sencilla.
- **Criterios subjetivos:** Son más difíciles de evaluar porque tratamos de medir la **satisfacción** que le ha proporcionado al usuario utilizar el producto.



Ilustración 2. Factores clave en usabilidad.

6.1.3. Principios de usabilidad

Los principios de usabilidad nos permiten estandarizar las soluciones que el diseño del producto requiere para que los usuarios puedan interactuar con él. El diseñador es libre de aplicarlos en función del contexto del producto. Así se distinguen unos principios básicos y generales de usabilidad.

Principios básicos

- **Facilidad de aprendizaje.** Se centra en los nuevos usuarios y cómo utilizan el producto.

- **Facilidad de uso.** Analizamos si los usuarios se manejan de forma natural y si usan más o menos pasos de los necesarios.
- **Flexibilidad.** Hace referencia a la diversidad con la que un usuario y el producto intercambian información.
- **Robustez.** Mide el soporte que adquiere el usuario por parte del producto a la hora de interactuar con él y finalizar sus metas.

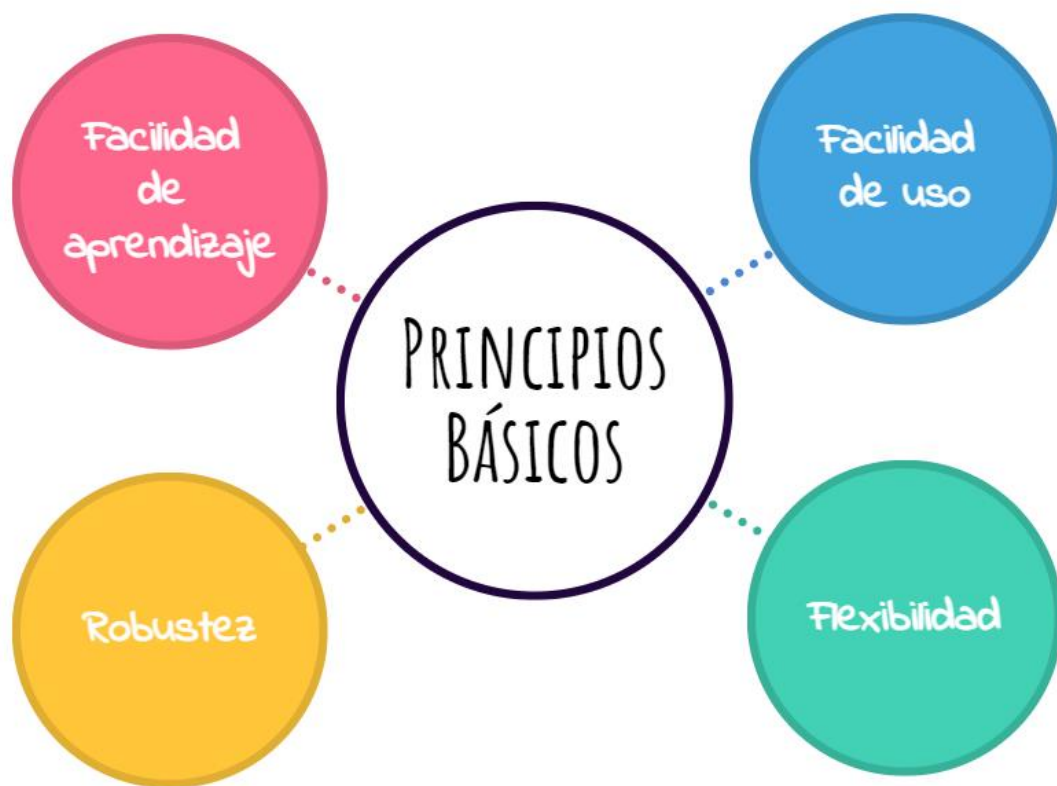


Ilustración 3. Principios básicos de usabilidad

Principios generales

- **Empatía con los usuarios.** Debemos conocer muy bien a nuestro usuario para diseñar el producto de forma que funcione tal y como espera.
- **Documentación y material de apoyo.** Debemos poner a disposición del usuario la ayuda, guías, tutoriales, documentación y cualquier otro material de apoyo que faciliten su aprendizaje.

- **Prevención de errores y retroalimentación.** El producto no sólo debe predecir los posibles errores que puedan cometer los usuarios, sino que además debe ayudarles a recuperarse de ellos mediante mensajes sencillos.
- **Facilidad de aprendizaje y uso.** Si conocemos al usuario, seremos capaces de diseñar el producto acorde a su experiencia y si además utilizamos un lenguaje sencillo y seguimos los estándares de la industria reduciremos en gran medida su curva de aprendizaje.
- **Flexibilidad.** Debemos ofrecer al usuario múltiples formas de comunicarse con el producto y distintas vías de personalización para que sientan que en todo momento tienen el control.
- **Consistencia.** El producto debe resultar familiar al usuario y comportarse de la misma forma independientemente del dispositivo que esté usando.
- **Robustez.** Es el grado de apoyo que recibe el usuario por parte del producto para ejecutar tareas y finalizar metas.
- **Adecuación.** Los usuarios tienen que poder realizar las tareas escogiendo la forma que quieran independientemente del número de pasos que necesiten o su complejidad, lo más importante es que puedan finalizarlas en el tiempo más rápido posible.
- **Tiempos de respuesta.** El producto debe tener un alto grado de respuesta para que el usuario mantenga la sensación de que todo fluye de forma natural y no pierda el foco.
- **Disminución de la carga cognitiva y accesibilidad.** Debemos tener presente a los usuarios con algún tipo de discapacidad y tratar de que tengan que recordar antes que memorizar cuando usan el producto.



Ilustración 4. Principios generales de usabilidad

6.1.4. Principios heurísticos de diseño de interacción de Tognazzini (Tognazzini, 2014)

Aunque las heurísticas de *Nielsen* (Nielsen, 10 Usability Heuristics for User Interface Design, 1995) son las más utilizadas y extendidas a la hora de evaluar la usabilidad de un producto, se han elegido para este marco teórico los principios heurísticos de diseño de interacción de *Tognazzini* por ser aquellos que se adaptan de forma más adecuada a la construcción de sitios web y que profundizan más en la interacción del usuario con la interfaz.

Estos principios están orientados al diseño e implementación de interfaces gráficas en las que los usuarios puedan comprender rápidamente cómo utilizar el producto diseñado y que a su vez le ayude a lograr sus objetivos en la interacción con el sitio web:

- **Estética.** El diseño visual del sitio web es casi tan importante como el diseño de usabilidad incluso teniendo un buen comportamiento del sitio una mejora en la estética aportará una mejora en su percepción.
- **Anticipación.** Se deben analizar las tareas en profundidad para anticiparnos a las necesidades que los usuarios puedan tener para resolverlas y proporcionarles todas las herramientas e información necesaria para cada paso del proceso.
- **Autonomía.** El sitio web debe proporcionar al usuario el control, permitir que tome sus propias decisiones aún en detrimento de empeorar la estética.
- **Color.** Se debe utilizar el color alejándose de modas pasajeras de forma inteligente que permitan transmitir información y teniendo en cuenta a los usuarios con alguna discapacidad visual.
- **Consistencia.** Hay que dar al sitio web una consistencia para que el usuario no se cree falsas expectativas ya que en ocasiones podemos cometer el error de implementar una funcionalidad de tal manera que el usuario espere que vaya a llevar a cabo una acción distinta a la que fue implementada.
- **Valores por defecto o predeterminados.** Los valores por defecto facilitan la interacción con el producto, pero no todo tiene porqué tener valores por defecto y estos deben ser fácilmente reemplazados.
- **Eficacia de usuario.** Debemos maximizar la productividad del usuario cuando utilice el producto y no tratar de buscar sólo la eficiencia de este.

- **Interfaces explorables.** El usuario debe poder regresar al comienzo en cualquier momento, permitir que puedan deshacer sus acciones, así como ofrecer distintas rutas que permitan a los usuarios más experimentados ejecutar las tareas directamente mientras que los usuarios noveles pueden experimentar libremente y proporcionar a todos una salida. A la hora de diseñar los elementos en un sitio web se deben tener en cuenta las leyes de *Fitts* y *Hick*; serán más fáciles de seleccionar por parte del usuario aquellos elementos que tengan un tamaño adecuado y no requieran un gran desplazamiento del cursor.
- **Objetos humanos.** Los objetos del sitio web deben seguir los estándares de esta forma aseguramos su familiaridad y consistencia. Por ejemplo, si incluimos un botón en el sitio web se espera que este pueda ser presionado.
- **Reducción de latencia.** Debemos implementar señales de percepción de la latencia tales como una barra de progreso cuando se ejecutan acciones que requieren un tiempo de espera, mensajes que muestren la duración estimada de una acción o visualizaciones de iconos de espera cuya duración sea de hasta 2 segundos.
- **Aprendizaje.** El mejor escenario es aquel en el que el usuario puede utilizar el sitio web desde el principio reduciendo al máximo la curva de aprendizaje.
- **Uso de metáforas.** Se han de elegir metáforas muy extendidas y fácilmente reconocibles que permitan a los usuarios captar rápidamente el concepto que se quiere transmitir.
- **Protege el trabajo del usuario.** Debemos asegurarnos de que el usuario nunca va a perder su trabajo ni por un error suyo ni por un error del sitio web.
- **Legibilidad.** Tenemos que elegir fuentes tipográficas que faciliten la lectura, seleccionar un tamaño de fuente y un contraste adecuado pensando en todo tipo de usuarios incluyendo aquellos con alguna discapacidad visual.

- **Estado.** Debemos ofrecer al usuario opciones de personalización para que el sitio web se adecue a la experiencia que ya tiene usándolo y toda esta información debe almacenarse de forma segura.
- **Navegación visible.** El usuario debe tener siempre la sensación de control y seguridad sobre el sitio web procurándole la percepción de que se encuentra siempre en el mismo entorno.



Ilustración 5. Principios heurísticos de diseño de interacción de Tognazzini

6.1.5. *Diseño Centrado en el Usuario (DCU)*

Es una metodología que sitúa al usuario en el centro y que tiene una secuenciación iterativa en la que el diseño se prueba y se optimiza hasta que alcanza el nivel de calidad deseado. Consta de las siguientes fases:

1. **Análisis.** Se establece el concepto del producto en base a la investigación de los usuarios y la comparación con otros productos existentes. En esta fase se marcan los objetivos del producto, quiénes son nuestros usuarios y qué requisitos técnicos va a necesitar el producto.
2. **Modelado de usuario.** Se trata de definir el perfil del usuario, para ello se suele utilizar el modelo *Persona* que puede incluir detalles como su perfil geográfico, demográfico y psicosocial, su relación con el producto, si es un cliente o un usuario ocasional y también podemos definir lo que se denominan *Scenarios* que recrean casos concretos de utilización del producto.
3. **Diseño.** Que consta de las siguientes etapas:
 - **Arquitectura de la información.** Es “el arte, la ciencia y la práctica de diseñar espacios interactivos comprensibles, que ofrezcan una experiencia de uso satisfactoria facilitando el encuentro entre las necesidades de los usuarios y los contenidos y/o funcionalidades del producto.” (Montero, 2015)
 - **Diseño conceptual.** En esta fase definimos el esquema de organización y navegación entre los contenidos. Establece las relaciones entre las diferentes páginas.
 - **Diseño visual.** Se definen las características gráficas de la interfaz en base a las fases anteriores. Aquí se realizan distintas acciones como la creación de un libro de estilo, documentación gráfica, diseño de la retícula, elección de la paleta de colores y la tipografía, etc.
 - **Diseño de contenidos.** Las principales características a tener en cuenta son la brevedad, facilitar la lectura en diagonal, utilizar lenguaje estructurado, legibilidad, títulos de página, tamaño de la tipografía, contraste entre texto y fondo, etc.

2. Prototipo

- **Baja fidelidad o *wireframe*.** Es una propuesta esquematizada de la estructura principal.
- **Alta fidelidad.** Tiene un aspecto muy parecido al diseño final.

3. **Evaluación.** Es la fase más importante del DCU y puede llevarse a cabo en varias etapas durante el desarrollo del producto y también en su etapa final. Pueden llevarse a cabo con usuarios reales o mediante metodología heurística.



Ilustración 6. Diseño centrado en el usuario (DCU)

7.1.6. Identificación de métricas cuantitativas (Fernández Casado, 2021)

Las métricas cuantitativas permiten obtener una medida objetiva y se utilizan para la toma de decisiones en cuanto al diseño del producto. Las principales métricas en usabilidad son:

- **Eficacia de la eliminación de defectos (Métrica de eficiencia).** Mide los errores encontrados durante el desarrollo frente a los de la fase final. Cuanto más cercano sea este valor a cero mayor será la eficiencia eliminando defectos.

$$EED \text{ (eficacia eliminación de defectos)} = \frac{E}{E + D}$$

E, número de errores encontrados en la fase de desarrollo.

D, número de errores encontrados en la entrega final.

- **Disponibilidad de la interfaz o sistema (Métrica de eficiencia).** Probabilidad de que el producto funcione bien dados unos requerimientos en un momento dado. Se utiliza en las fases de mantenimiento del producto y cuanto menor sea el tiempo de recuperación mayor será su disponibilidad.

$$D \text{ (disponibilidad)} = \frac{T_{MDF}}{T_{MDF} + T_{MDR}} \times 100$$

T_{MDF} , tiempo medio por fallo.

T_{MDR} , tiempo medio de recuperación.

- **Fiabilidad de la interfaz o sistema (Métrica de eficiencia).** Es el tiempo que los usuarios utilizan el producto sin errores. Se suele emplear en la fase de desarrollo y cuanto mayor sea su valor mayor será la fiabilidad.

$$T_{MEF} = T_{MDF} + T_{MDR}$$

T_{MEF} , tiempo medio entre fallos.

T_{MDF} , tiempo medio por fallo.

T_{MDR} , tiempo medio de recuperación.

- **Número de errores (Métrica de efectividad).** Se recoge cualquier tipo de error, ya sea involuntario o no por parte del usuario mientras ejecutaba alguna tarea con el producto. Los errores se clasifican junto con una descripción y un grado de importancia y dan una visión de la calidad del producto.
- **Número de errores por unidad de tiempo (Métrica de efectividad).** Es la métrica del número de errores, pero en un periodo de tiempo concreto y se usa en etapas de desarrollo y mantenimiento.
- **Número de tareas que pueden realizarse (Métrica de efectividad).** A partir de esta métrica podemos saber el grado de complejidad del producto.
- **Porcentaje de enlaces rotos (Métrica de eficiencia).** Es importante desde el punto de vista de la usabilidad y accesibilidad ya que un enlace roto puede suponer que el usuario no termine una tarea. Cuanto menor sea el porcentaje, mayor será la satisfacción del usuario ya que el producto resulta más usable y accesible.

$$ER\% (\% \text{ de enlaces rotos}) = \frac{ER_I + ER_E}{ER} \times 100$$

ER_I , número de enlaces rotos internos.

ER_E , número de enlaces rotos externos.

ER , número total de enlaces rotos.

- **Porcentaje de presencia de propiedades en los objetos o elementos (Métrica de eficiencia).** Proporciona el porcentaje de elementos que carecen de una propiedad concreta, suele utilizarse para medir la aparición o no de etiquetas en el código como ALT o TITLE que sirven de ayuda a los usuarios con dificultades visuales. Cuanto menor sea este porcentaje más usable y accesible es el producto y por tanto mayor satisfacción es la del usuario.

$$P_{PC} (\% \text{ de presencia}) = \left(\frac{E_{SP}}{E} \right) \times 100$$

E_{SP} , número de elementos sin la propiedad a medir.

E , número de elementos totales.

- **Porcentaje de tareas finalizadas con éxito (Métrica de efectividad).** Mide la ratio de tareas finalizadas con éxito independientemente de si se han producido errores o no durante su ejecución. Se utiliza un sistema binario, asignando un valor de 1 si la tarea ha tenido éxito y 0 si no. Proporciona una idea de la facilidad de uso del producto.

$$P_{TE} (\% \text{ de tareas finalizadas con éxito}) = \left(\frac{T_E}{T_E + T_F} \right) \times 100$$

T_E , número de tareas finalizadas con éxito.

T_F , número de tareas no superadas.

- **Porcentaje de tareas finalizadas con éxito en el primer intento (Métrica de efectividad).** De nuevo empleamos el sistema binario y nos da una idea de facilidad de uso.

$$P_{TE1} (\% \text{ de tareas finalizadas con éxito en el primer intento}) = \left(\frac{T_{E1}}{T_{E1} + T_F} \right) \times 100$$

T_{E1} , número de tareas finalizadas con éxito en un primer intento.

T_F , número de tareas no superadas.

- **Porcentaje de usuarios capaces de finalizar las tareas clave sin ayuda (Métrica de efectividad).** Cuanto más cerca esté del valor del porcentaje de tareas finalizadas con éxito mayor será la usabilidad.
- **Tiempo de permanencia (Métrica de eficiencia).** Tiempo que ha permanecido un usuario utilizando el producto.
- **Tasa de uso (Métrica de eficiencia).** Se basa en tres principios:
 - **Facilidad de aprendizaje.** Tiempo que tarda un usuario nuevo en realizar una tarea.
 - **Eficiencia de uso.** Tiempo que tarda un usuario recurrente en realizar la misma tarea que realiza un usuario nuevo.
 - **Recuerdo en el tiempo.** Tiempo que tarda un usuario recurrente en realizar la misma tarea pasado un tiempo.

La forma de medirla es a través de la media y la desviación estándar sobre los tiempos invertidos:

$$M = \frac{(FA + EU + RT)}{3}$$
$$\sigma = \sqrt{\frac{(FA - M)^2 + (EU - M)^2 + (RT - M)^2}{2}}$$

Siendo:

FA, tiempo de facilidad de aprendizaje.

EU, tiempo de la eficiencia de uso.

RT, tiempo de recuerdo en el tiempo.

M, media aritmética de los tiempos FA, EU y RT.

σ , desviación estándar.

A valores altos de desviación estándar la usabilidad de la tarea es baja y por tanto debería volver a diseñarse para disminuir este valor.

- **Tiempo de latencia (Métrica de eficiencia).** Tiempo que tarda el producto en responder y que se asocia al tiempo que un usuario puede mantener su atención.
- **Click Trough Rate (CTR) (Métrica de eficiencia).** Mide el número de veces que se ha mostrado una impresión frente al número de veces que se ha hecho clic. A mayor CTR mayor efectividad tendrá sobre los usuarios.

$$CTR = \frac{\text{Número de visualizaciones}}{\text{Número de clics}}$$

- **Páginas vistas / clics.** Existe una alta correlación entre los clics y el tiempo dedicado a cada tarea.
- **Tasa de conversión (Métrica de efectividad y satisfacción).** Indica el porcentaje de usuarios que han finalizado una tarea marcada como un objetivo. Se utiliza el sistema binario donde 1 es que se ha llevado a cabo una conversión y 0 es que no.

$$CR = \left(\frac{O_c}{T_v}\right) \times 100$$

Siendo:

CR, tasa de conversión.

O_c , es el número de objetivos conseguidos.

T_v , es el número total de visitas.

La media de conversión varía de un producto a otro, aunque se fija entre un 1 y un 3 por ciento. Cuando esta ratio de conversión es baja puede ser consecuencia de distintos factores entre ellos problemas de usabilidad consecuencia de un mal diseño.

- **Tasa de rebote (Métrica de efectividad y satisfacción).** Mide el porcentaje de visitantes que abandona el producto sin interactuar con él. De nuevo y entre diversos factores un mal diseño que provoca problemas de usabilidad puede ser una causa por la que el usuario abandona. El tiempo estimado para descartar un rebote se ha fijado en 30 segundos, más allá de ese tiempo se considera que el usuario está interesado y por tanto no es un rebote.

$$T_R (\text{tasa de rebote}) = \frac{T_U}{T_V}$$

T_U , número total de visitas que acceden una única vez.

T_V , número total de visitas.

- **Ticket medio.** Es una métrica utilizada en los *ecommerce* para calcular la media de lo que gastan los usuarios.

$$T_M (\text{ticket medio}) = \frac{IT}{NP}$$

IT , ingresos totales.

NP , número de pedidos.

6.2 SEO

6.2.1. Definición

El SEO es el conjunto de acciones que tratan de mejorar el posicionamiento de un sitio web en un buscador, normalmente *Google* por ser el de uso mayoritario.

Pero el SEO va mucho más allá “parte del conocimiento de dos actores principales: el funcionamiento del propio buscador y el comportamiento de quienes recurren a él para solventar su necesidad”. (Maciá Domene, 2019)

6.2.2. ¿Cómo funciona un buscador?

Un **buscador** es un sistema que almacena grandes cantidades de información que han sido almacenadas previamente en un índice a partir de sus contenidos como imágenes, audios, videos, referenciados a los documentos rastreados y que es capaz de recuperarlos en función de una búsqueda realizada por el usuario.

Este proceso de recuperación de la información se activa cuando un usuario **expresa su necesidad** introduciendo una palabra o frase en el buscador. Hay consultas que tienen una **respuesta directa** y otras que pueden generar respuestas en **múltiples resultados** que son ordenados por relevancia.

Para que un contenido aparezca en las búsquedas debe ser **indexable**, es decir, que ha de cumplir una serie de condiciones como que el contenido sea **público, accesible, no esté bloqueado y no necesite de la interacción humana** para poder ser rastreado.

El **proceso de indexación** de contenidos consta de las siguientes fases:

1. **Descubrimiento.** Para que un buscador añada contenido éste ha de ser descubierto. *Google* visita con distinta frecuencia los sitios que ya han sido indexados en busca de nuevo contenido. En el caso de que sea un nuevo sitio web podemos comunicar su existencia a través de *Google Search Console* incluyendo un archivo *sitemap*.

- 2. Rastreo.** Cuando un buscador descubre una nueva página el robot rastrea todo el contenido para clasificarlo; a partir de ese momento visitará la página cada cierto tiempo para actualizar el contenido. El sitio web deberá estar correctamente optimizado para aprovechar la cuota de rastreo (*crawl-budget*) que el robot emplea. Cuanto más optimizado, más fácil le resultará rastrear el sitio e indexar más contenido.
- 3. Renderización.** *Google* rastrea no solo los contenidos, sino también el CSS, JS, fuentes tipográficas, imágenes, etc. para tratar de interpretar toda esa información como si fuera un navegador que es lo que en definitiva ve un usuario.
- 4. Clasificación.** *Google* emplea la renderización y toda la información rastreada junto con el marcado semántico, marcado de datos estructurados y etiquetas HTML para clasificar el sitio web dentro de una determinada categoría de búsquedas.
- 5. Indexación.** El siguiente paso es que *Google* añadirá una referencia del sitio web a su índice y a partir de ese momento éste podría aparecer como resultado de una búsqueda.
- 6. Posicionamiento.** Para cada página rastreada el buscador asigna una posición con respecto al resto de páginas indexadas en función de su relevancia, popularidad, frescura de contenido y cómo los usuarios interactúan con ella.



Ilustración 7. Fases del proceso de indexación

6.2.3. Intención de búsqueda

Una búsqueda nace con una necesidad que el usuario manifiesta y puede ser de distintos tipos:

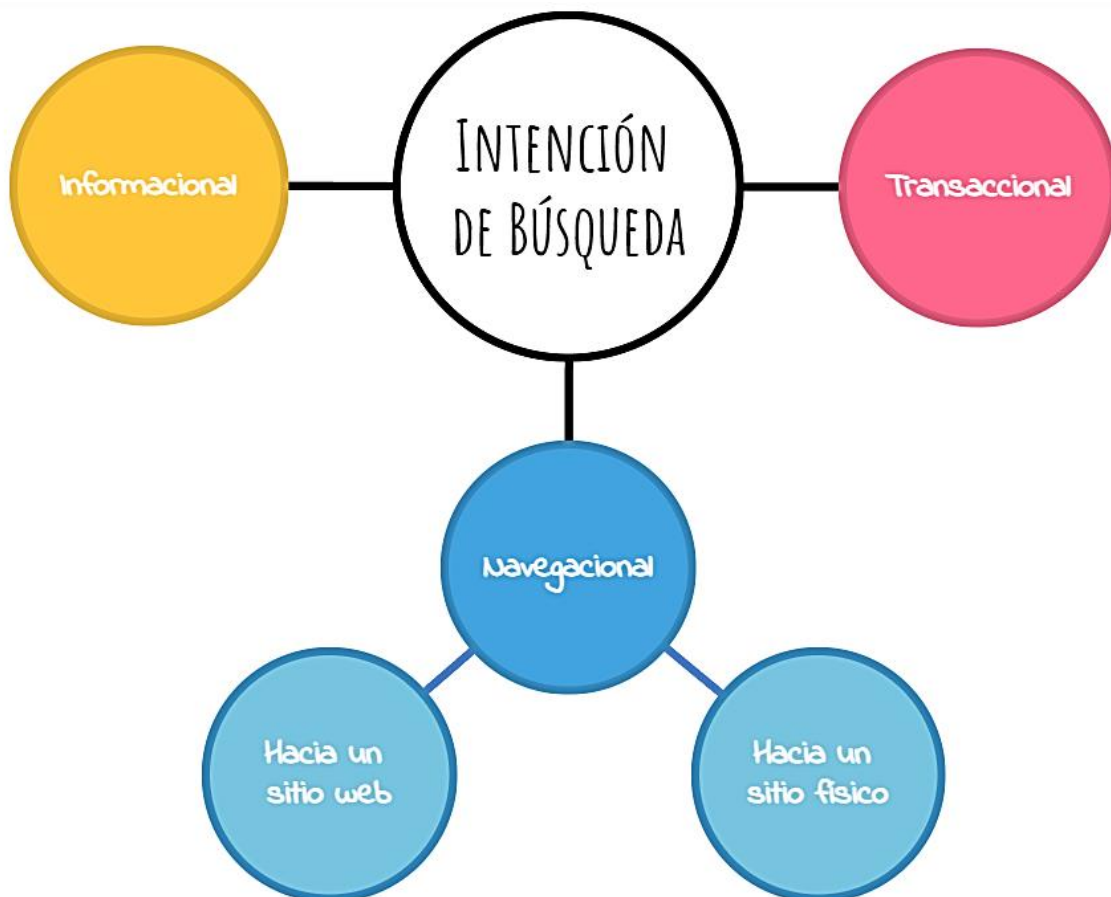


Ilustración 8. Intención de búsqueda,

- **Informacional (HOW).** Necesidad de saber. *Google* es capaz de distinguir entre búsquedas con una respuesta directa y búsquedas más amplias.
- **Transaccional (DO).** Necesidad de realizar una acción.
- **Navegacional hacia un sitio web (VISIT WEBSITE).** Los usuarios quieren visitar un sitio web, pero no recuerdan el dominio.
- **Navegacional hacia un sitio físico (VISIT-IN-PERSON).** Búsquedas que se realizan desde dispositivos móviles, el usuario está buscando un lugar físico.

6.2.4. El perfil del cliente objetivo: *Searcher Persona*

En otros ámbitos, como en el DCU (Diseño centrado en el usuario), se emplea la construcción de un cliente imaginario que combina los datos de un perfil de cliente con su comportamiento. Trayendo este concepto al SEO podemos definir el *searcher persona* donde se incluyen los siguientes apartados:

- **Nombre.** Para establecer una relación cercana con el usuario se introduce un nombre y una imagen.
- **Trabajo y cargo.** Se especifica si su ámbito profesional va a estar relacionado con las intenciones de búsqueda que pueda tener o no. Este factor es fundamental en la selección de palabras clave, ya que un usuario común utilizará un lenguaje más coloquial, mientras que un usuario profesional realizará un tipo de búsqueda más técnica.
- **Datos demográficos.** Datos relativos a la edad, sexo, capacidad adquisitiva, localización geográfica, nivel educativo, tamaño de la familia, etc.
- **Objetivos y retos.** En este apartado se reflejan cuáles son las principales metas para alcanzar por el usuario y a qué obstáculos se enfrenta referidos en todo momento al sitio web. De este apartado podemos extraer la información necesaria para construir el *title* y la *meta description* de las páginas del sitio, ya que nos dan una idea de cuáles son las motivaciones de la *searcher persona* para tratar de llamar su atención en los resultados del buscador, así como hacer un marcado de datos estructurados adecuado a esta información.
- **Valores y miedos.** Hacen referencia a las motivaciones y objeciones a la hora de finalizar un objetivo en el sitio web entre las que debemos tener en cuenta aquellas que pueden retrasar la decisión o incluso paralizarla. Detectar estos puntos de fricción es importante para anticiparnos y ayudar a los usuarios a superarlos. Este apartado da las claves acerca de la estructura semántica y el contenido que ha de seguir la página ya que podremos reflejar los mejores argumentos para finalizar el objetivo y anticipar respuestas a las principales objeciones del usuario.

- **Propuesta de valor.** Es un resumen de todos los puntos anteriores.



Ilustración 9. Searcher persona.

6.2.5. El viaje del cliente o customer journey

Comprende todo cuanto hace un usuario desde que es consciente de su necesidad hasta que toma la decisión final de realizar una compra o finalizar un objetivo en el sitio web.

El viaje da comienzo cuando el usuario intuye su necesidad y realiza una consulta en el buscador; entonces se produce la primera impresión del sitio web en los resultados. Una primera posición orgánica, es decir, aquella que aparece en primera posición después de los anunciantes puede llegar a recibir entre un veinte y un cuarenta por ciento de clics.

En esta fase el usuario quizá no tenga clara aún su necesidad así que el objetivo es intentar aparecer en búsquedas genéricas relacionadas con los intereses de la *searcher persona*, es decir, búsquedas informacionales o KNOW.

A continuación, el usuario ha concretado su necesidad y empieza a valorar distintas alternativas para elegir aquella que crea que cumple con sus expectativas y que supera sus puntos de fricción. Estas búsquedas se corresponden con las búsquedas transaccionales o HOW.

Finalmente, el usuario ha tomado una decisión, así que su intención de búsqueda pasa a ser transaccional o DO. En esta fase convivirán las búsquedas navegacionales *visit website* y *visit in person*.



Ilustración 10. Viaje del cliente.

6.2.6. *Keyword Research*

El siguiente paso en una estrategia de posicionamiento es el de la creación de contenidos. El viaje del cliente y la *searcher persona* creados son los puntos clave de partida y además se han de tener en cuenta en la elección de palabras clave, los aspectos de **profundidad** y **amplitud**, ya que dependiendo uno y otro nos llevará a estrategias de posicionamiento distintas:

- **Profundidad.** Hace referencia a la visibilidad en el buscador en función de la posición y del tipo de resultado para una palabra clave que para una primera posición del bloque orgánico pueden ser: fragmentos destacados, respuestas directas, fragmentos enriquecidos, resultados locales, carrusel de imágenes, vídeos, etc.
- **Amplitud.** Es la visibilidad alcanzada a partir del posicionamiento en un mayor número de palabras clave relacionadas distintas lo que también incrementa la presencia en el buscador.

6.2.7. *Metodología SEO*

La principal dificultad a la hora de aplicar una metodología SEO es que no tiene unos estándares, cada agencia o profesional puede aplicar una metodología distinta con mayor o menor acierto en base a su propia experiencia; además el algoritmo de *Google* es cambiante y continuamente obliga a los profesionales a lanzar hipótesis, realizar distintas pruebas y tratar de medir los resultados para contrastarlas.

Las metodologías SEO deben aplicarse a un sitio web de nueva creación desde el inicio; también se pueden aplicar a sitios webs maduros, aunque sea más difícil de implementar y conlleve algunos riesgos.

Se ha elegido la metodología planteada por Fernando Maciá en su libro sobre SEO Avanzado (Maciá Domene, 2019) que bajo su experiencia plantea cinco grandes bloques que contienen una serie de puntos a validar en cada uno de ellos.



Ilustración 11. Metodología SEO.

- 1. Indexabilidad.** Para que un contenido de un sitio web aparezca en los resultados de búsqueda debe ser **descubierto**, **rastreado**, **clasificado** y finalmente **indexado**.

Este punto es un factor importante y que ha de cumplirse sobre todo para aquellas páginas del sitio web que sean más **relevantes**, que tengan un objetivo para el usuario o cumplan con una intención de búsqueda.

Igual de importante es evitar que páginas con contenido **irrelevante** (*thin content*) o **duplicado** sean indexadas por el buscador.

Para ello se debe asegurar que técnicamente se cumplen los requisitos en el sitio web tanto para indexar páginas como para impedir su indexación.

En esta parte se realizan acciones como:

- Conectar el dominio en *Google Search Console*.
- Verificar cuántas páginas del sitio han sido indexadas.
- Crear el archivo *sitemap*.
- Comprobar el contenido duplicado y uso del enlace *canonical*.
- Conservar en todo momento la raíz del dominio.
- Gestión de contenidos paginados.
- Configurar la *meta robots* con la que se bloquea la indexación o no e indicar si debe continuar para encontrar nuevos contenidos o no.
- Revisar la caché de *Google*.
- Vista de página con *JavaScript* desactivado con la que detectar problemas que impiden la navegación.
- Vista de página con *cookies* desactivadas con la que revelar problemas relacionados con la personalización del usuario.
- URL que reciben variables pasadas por el método POST.
- Comprobar el orden de carga de la página para analizar si los contenidos relevantes para el posicionamiento son los primeros en descargar.
- Renderizado de página.
- Crear el archivo *robots.txt* y sus directivas.
- Atributos *rel=nofollow*, *ugc* o *sponsored* en enlaces del sitio.
- Creación de enlaces amigables.

2. Contenido. La calidad del contenido es otra parte fundamental en una estrategia SEO y aunque es algo subjetivo y difícil de analizar se debe buscar en todo momento aportar valor cumpliendo las expectativas que el usuario tiene cuando expresa su necesidad a través de la intención de búsqueda que ha generado.

Si el usuario se encuentra satisfecho con el contenido que encuentra realizará acciones positivas como escribir comentarios, compartir en redes sociales, recomendar, permanecer tiempo en la página e incluso generar enlaces al sitio web de forma natural.

En cuanto a la implementación técnica se comprueba si el gestor de contenidos o tecnologías que estamos usando para construir el sitio web facilitan la implementación SEO de las principales áreas del contenido cómo títulos, encabezamientos, textos alternativos o marcado de datos estructurados.

Se realizan acciones como:

- Optimizar los títulos en *SERPs (Search Engine Result Pages)*.
- Optimizar el *Snippet* en SERPs.
- Uso de *keywords* de forma natural.
- Uso correcto de las marcas de jerarquía HN.
- Revisión del peso y tamaño de las imágenes y sus etiquetas.
- Enfoque del contenido de la página ya que cuanto más específico sea el tema por tratar mejor será categorizado por *Google*.
- Evitar incluir contenido en tecnologías *Flash/Silverlight*.
- Evitar incluir contenido en *frames* o *iframes*.
- Uso de *AJAX (Asynchronous Javascript and XML)*.
- Análisis de cabeceras HTTP, donde se analizan los códigos de respuesta del servidor 200, 301, 302, 404, 410.
- Revisión de la inclusión de IP en listas negras.
- Revisión del vecindario de IP en el caso de estar en un hosting compartido.
- Antigüedad del dominio.
- Estabilidad de URL en el tiempo.
- Revisar la existencia del protocolo HTTP/HTTPS.

3. *Mobile first.* La compatibilidad entre los distintos dispositivos desde los que puede acceder un usuario es un factor clave en cuanto al posicionamiento, de hecho, *Google* ya presenta sus resultados teniendo en cuenta la navegación móvil por ser esta la forma más extendida en la que los usuarios hacen uso del buscador.

Para ello se ha de implementar un diseño adaptativo, es decir, que el sitio web se ajuste a los distintos formatos de pantalla utilizando el mismo código HTML y URL para cada página.

4. WPO (Web performance Optimization). La optimización en el rendimiento de un sitio web es un factor clave en el posicionamiento, además la mayor parte del tráfico proviene de *smartphones* cuyo ancho de banda es más limitado.

Se revisan los siguientes puntos:

- Uso de HTTPS con HTTP/2.
- Tiempo de descarga del código HTML.
- Tiempo de carga total en la que la página está completamente operativa.
- Tiempo de resolución del DNS (*Domain Name System*).
- Tiempo de conexión que es el tiempo invertido por el navegador en establecer la conexión con el servidor.
- Tamaño de archivos de página. Cuanto más ligera, mejor.
- Tiempos orientados al usuario. *Google* también tiene en cuenta los tiempos de espera desde la perspectiva del usuario.
 - **FP (First Paint).** Ocurre cuando se pinta el primer píxel de la página.
 - **FCP (First Contentful Paint).** *Google* recopila este dato de los propios usuarios que navegan con *Chrome* y afecta al posicionamiento de la página. Suele coincidir con el valor de FP.
 - **FID (First Input Delay).** Tiempo en que se retrasa la primera interacción del usuario.
 - **Speed Index.** Es un indicador del nivel de eficacia con que nuestra web pinta el contenido en el *above the fold* o parte visible de la web. Es muy similar el FMP (*First Meaningful Paint*)
- JavaScript y CSS en archivos externos.

5. Perfil de Popularidad. Consiste en una revisión de los enlaces entrantes y los dominios a los que pertenecen estos enlaces para verificar su procedencia y asegurar su naturalidad o toxicidad.

Se realizan las siguientes acciones:

- Enlaces externos entrantes. Analizamos el número y calidad de estos enlaces ya que supone uno de los elementos de relevancia más importante en el posicionamiento.
- Dominios satélites pertenecientes a la misma empresa como un blog o un sitio web creado para una campaña concreta.
- *Anchor text* en los enlaces entrantes.
- Dominios entrantes sospechosos.

6.2.8. Identificación de métricas cuantitativas

Los indicadores clave de rendimiento KPI de un proyecto SEO son:

1. Indicadores relacionados con la profundidad del posicionamiento. Miden cómo evoluciona el posicionamiento del sitio web para las palabras clave más importantes.

Se utilizan los siguientes indicadores:

- **Posiciones para nuestra propia selección de *keywords*.** Se analiza el posicionamiento para la lista de palabras clave creada para el sitio web incluyendo a los competidores.
- **Distribución de las posiciones de *keywords*.** Aquellos sitios web relevantes en su sector tendrán la mayor parte de las búsquedas ubicadas en primera página, sin embargo, aquellos menos relevantes las tendrán a partir de la segunda página.
- **Distribución de las posiciones por URL.** Se analiza el número de páginas del sitio web que ocupan posiciones en primera página.
- **Posicionamiento en resultados destacados.** Son resultados para contenidos que responden a una respuesta directa en las búsquedas.

2. Indicadores relacionados con la amplitud del posicionamiento. Se analizan todas aquellas palabras clave relacionadas con una intención de búsqueda transaccional. Se utilizan los siguientes indicadores:

- **Páginas indexadas totales o saturación.** Relación entre el número de páginas indexadas y el número de páginas que contiene el sitio web. El objetivo es indexar el mayor número de páginas posible.

$$\text{Saturación} = \frac{\text{número de páginas indexadas}}{\text{número de páginas del sitio web}}$$

- **Saturación por debajo del 100%.** Problemas de accesibilidad derivado del uso de tecnologías que impiden el rastreo o hay problemas con la calidad del contenido. También puede haber una errónea codificación del archivo robots.txt
 - **Saturación cerca del 100%.** Situación óptima.
 - **Saturación muy por encima del 100%.** El sitio web tiene problemas de contenido duplicado porque se ha indexado contenido con www y sin www o con el protocolo http y https a la vez o porque se ha indexado desde un servidor utilizado en preproducción.
- **Páginas indexadas por sección.** Cantidad de páginas del sitio web indexada por temática, categoría o sección.
 - **Número de URL únicas posicionadas.** Análisis de cuántas URL están posicionadas para alguna búsqueda en el sentido más amplio y no sólo reducido a nuestra lista de palabras clave.
 - **Número de búsquedas posicionadas.** Dato relacionado con la variedad de búsquedas donde contamos con páginas posicionadas, puede entenderse como la inversa de la anterior.
 - **Número de URL únicas como páginas de entrada de tráfico orgánico.** De todas las páginas indexadas y posicionadas del sitio web solo algunas recibirán tráfico orgánico así que se trata de determinar cuáles son y cuántas, lo que nos dará una idea de cuántas de ellas son realmente efectivas a la hora de atraer tráfico orgánico.

3. Indicadores relacionados con una mejora de la calidad del contenido.

- **Páginas excluidas en *Google Search Console*.** Si obviamos la parte técnica por la que una página puede no ser indexada, un contenido débil, duplicado o de baja calidad es motivo de exclusión del índice de *Google*.
- **Indicadores de calidad de la visita por página.** El contenido de las páginas debe satisfacer la intención de búsqueda del usuario. Aquí podemos medir la **tasa de rebote, las páginas vistas por sesión, la duración de la sesión, y la conversión.**

4. Indicadores relacionados con la mejora en la velocidad de descarga. En el caso en el que queramos optimizar el rendimiento del sitio web deberemos tener en cuenta los siguientes indicadores.

- **Estadística de rastreo.** Queremos saber el tiempo de descarga de una página y cuál es su tendencia en el tiempo.
- **Informe de velocidad en *Google Search Console*.** Analizaremos aquellas páginas del sitio web que tengan una velocidad baja, media o alta para todos los dispositivos y su tendencia a lo largo del tiempo.

5. Indicadores relacionados con oportunidades de mejora. Son métricas que apuntan a oportunidades de mejora rápida.

- **Canibalización de *keywords*.** Se identifican aquellas búsquedas para las cuales están compitiendo varias páginas en el sitio web.
- **URL posicionadas en segunda página.** Estas páginas son candidatas para mejorar su posicionamiento con un pequeño esfuerzo SEO.

6. Indicadores relacionados con una mejora del tráfico orgánico utilizando el segmento correspondiente en *Google Analytics* para analizar la cantidad total, evolución, procedencias, dispositivos de acceso, etc. para detectar problemas o caídas en el tráfico.

6.3 Combinación Usabilidad y SEO

Identificamos en este apartado aquellas áreas de la usabilidad y el SEO que han de ser trabajadas en común.

1. Análisis → Estudio de la competencia
2. Modelado de usuario → *Searcher persona, Customer Journey, Keyword Research*
3. Diseño
 - a. Arquitectura de la información → Indexabilidad
 - b. Diseño conceptual
 - c. Diseño de contenidos → Contenido
 - d. Diseño visual → *Mobile first*
4. Prototipo
 - a. Baja fidelidad o *wireframe*
 - b. Alta fidelidad → *WPO*
5. Evaluación

6.4 Herramientas de usabilidad

Principales herramientas para el análisis de usabilidad.

- **PageSpeed**, nos brinda un informe sencillo acerca de las principales métricas relacionadas con la experiencia de usuario relacionadas con el WPO.
- **Lighthouse**, es una extensión de Chrome que revisa la accesibilidad y el rendimiento del sitio web entre otras métricas.
- **Web Vitals**, es una extensión de Google para Chrome que mide la experiencia de usuario en cualquier sitio web. Métricas como la interactividad, la estabilidad visual y el tiempo de carga.
- **UX Check**, es una extensión de Chrome que sirve para facilitar las evaluaciones heurísticas.

- **Total Validator**, es una aplicación online que nos permite realizar distintas pruebas sobre nuestro sitio web. Revisa la accesibilidad, el HTML, el CSS, enlaces rotos y una revisión ortográfica.
- **Pingdom Tools**, es una herramienta online que nos da una idea del número total de objetos DOM, peso y velocidad de carga.
- **Legible**, es una aplicación online que nos permite analizar la legibilidad de los textos.
- **Hotjar**, es una herramienta con la que podemos monitorizar el comportamiento del usuario en el sitio web mediante la visualización de mapas de calor o sesiones grabadas completas.

6.5 Herramientas SEO

Recogemos en este apartado las principales herramientas a utilizar según la fase de metodología SEO en la que nos encontremos, así como aquellas relacionadas con las métricas.

Principales herramientas para Keyword Research

- **Para detectar búsquedas informacionales**, *Keywordtool.io*, *Google Suggest*, *Answerthepublic.com* y *SEMrush*.
- **Para extraer búsquedas a partir de nuestros competidores**. *Sistrix*, *SEMrush*, *Ubbersuggest* y *Similarweb*.

Principales herramientas para indexación

- **Google Search Console**, diagnosticar problemas de indexabilidad, nos informa de errores en el rastreo, problemas de cobertura, errores de interpretación de datos estructurados, creación de un archivo de desautorización de dominios, notificación de la aplicación de penalizaciones manuales, fallos de seguridad, contenido duplicado, podemos ver las páginas rastreadas como lo haría el robot de Google
- **Screaming Frog**, con ella se analiza el impacto de la etiqueta canonical en todo el sitio web.

- **Web Developer para Chrome**, que nos permite desactivar características relacionadas con *JavaScript* y *cookies* para detectar problemas de indexabilidad y redireccionamiento, comprobar el orden de carga de elementos en la página.

Principales herramientas para contenido

- **Web Developer para Chrome**, que nos permite analizar las respuestas del servidor.
- **Link Redirect Trace**, es una herramienta online que desglosa la respuesta del servidor a una URL determinada.
- **Google Search Console**, ha empezado a recoger recientemente unas métricas relacionadas con el contenido que más atrae a los usuarios denominadas *insights*.

Principales herramientas para mobile first

- **WebPageTest**, es una herramienta online que donde podemos elegir elegir un punto de partida y un navegador para medir la velocidad de descarga.
- **DNS Speed Test**, es una herramienta online que nos permite medir el tiempo de resolución DNS del dominio.
- **Web Developer para Chrome**, en su herramienta para desarrolladores podemos analizar todos los tiempos de carga y los tamaños de archivo del sitio web.
- **PageSpeed Insights** es un sitio web en el que tras introducir nuestro dominio podemos acceder a las principales métricas orientadas al usuario desde un móvil o desde una versión de escritorio.
- **Prueba de optimización para móviles** disponible en *Google Search Console* que nos permitirá analizar cómo es la representación gráfica del sitio web y detectar posibles errores.

Principales herramientas para el perfil de popularidad

- **Google Search Console** nos muestra un listado de enlaces que apuntan a nuestro dominio, queda en nuestra mano analizar la naturalidad y toxicidad de los mismos e implementar un archivo de desautorización en el caso de que detectemos enlaces tóxicos que puedan afectar a nuestro nivel de popularidad. También podemos revisar el anchor text de los enlaces entrantes.

Principales herramientas para KPI

- **Indicadores de la profundidad y amplitud del posicionamiento:** *SEMrush, Sistrix, Moz Pro, Ahrefs y Google Analytics.*
- **Indicadores para mejorar la calidad del contenido:** *Google Search Console y Google Analytics.*
- **Indicadores relacionados con la mejora en la velocidad de descarga.** *Google Search Console y PageSpeed Insights.*
- **Indicadores relacionados con oportunidades de mejora.** *Sistrix y Google Search Console.*
- **Indicadores relacionados con una mejora del tráfico orgánico.** *Google Analytics.*

7. Marco práctico

7.1 Datos de partida

Dado que *Google* contiene una alta cantidad de páginas web que ya superan los 100 millones de gigabytes, para elegir una muestra y realizar el siguiente análisis se ha partido de una base de datos de sitios webs cuyo tráfico está orientado al público español y que han sido agrupados por sectores y clasificados dentro de cada uno por orden de relevancia en *Google* después del último *Core Update* de Julio de 2021. (Cachón, 2021)

7.2 Selección del sector

El sector elegido para el estudio ha sido el de mascotas ya que la experiencia previa que tenemos en este sector puede ayudarnos a interpretar mejor las métricas y resultados que obtengamos en el análisis.

El sector recoge 27 sitios web con diferentes tipologías y que se recogen en la siguiente tabla:

Listado Sitios Web	Tipología
piensoymascotas.com	<i>Ecommerce</i>
mascotaplanet.com	<i>Ecommerce</i>
toppercan.es	Blog
todoperros.com	Blog
affinity-petcare.com	Marca
miscota.es	<i>Ecommerce</i>
soyunperro.com	Blog
mundoperros.es	Blog
hillspet.es	Marca
elcorteingles.es/mascotas	Ecommerce Generalista
tiendanimal.es	<i>Ecommerce</i>
purina.es	Marca
uncomo.mundodeportivo.com/animales	Medio Generalista
petclit.es	<i>Ecommerce</i>
demascotas.info	Blog
bekiamascotas.com	Blog
mascotasyamas.net	Blog

expertoanimal.com	Blog
depeces.com	Blog
kiwoko.com	Ecommerce
royalcanin.com	Marca
mundoanimalia.com	Blog
dogking.es	Blog
mascoteros.com	Ecommerce
nuestroperro.es	Ecommerce
razasyperros.es	Blog
bitiba.es	Ecommerce

Tabla 1. Listado sectorial Mascotas.

Debido a que las diferentes tipologías tienen distintos objetivos y con la intención de tomar una muestra lo suficientemente uniforme se eligieron aquellos sitios web orientados a la venta online y no generalistas. De manera que la muestra de partida fue la siguiente:

Listado Sitios Web
piensoymascotas.com
mascotaplanet.com
miscota.es
tiendanimal.es
petclic.es
kiwoko.com
mascoteros.com
nuestroperro.es
bitiba.es

Tabla 2. Listado ecommerce sector mascotas.

7.3 Estudio de competidores

Una vez seleccionada la muestra de partida se analizaron en primer lugar si realmente todos estos sitios web competían entre ellos; para ello se utilizó la herramienta *Ubersuggest*.

Es una aplicación online que ofrece información sobre los dominios a analizar como, por ejemplo: una visión general del dominio, las páginas que posicionan en el buscador, sugerencias de palabras clave, ideas de contenido y un listado de *backlinks*.

La principal ventaja de esta herramienta con respecto a otras es que utiliza la ingeniería inversa, es decir, que podemos obtener datos de los competidores para ver que

estrategias les están funcionando y aplicarlo a la nuestra, además de poder analizar nuestros propios datos.

Se introdujeron cada uno de los sitios seleccionados y se extrajeron para cada uno de ellos una lista de competidores en un archivo csv.

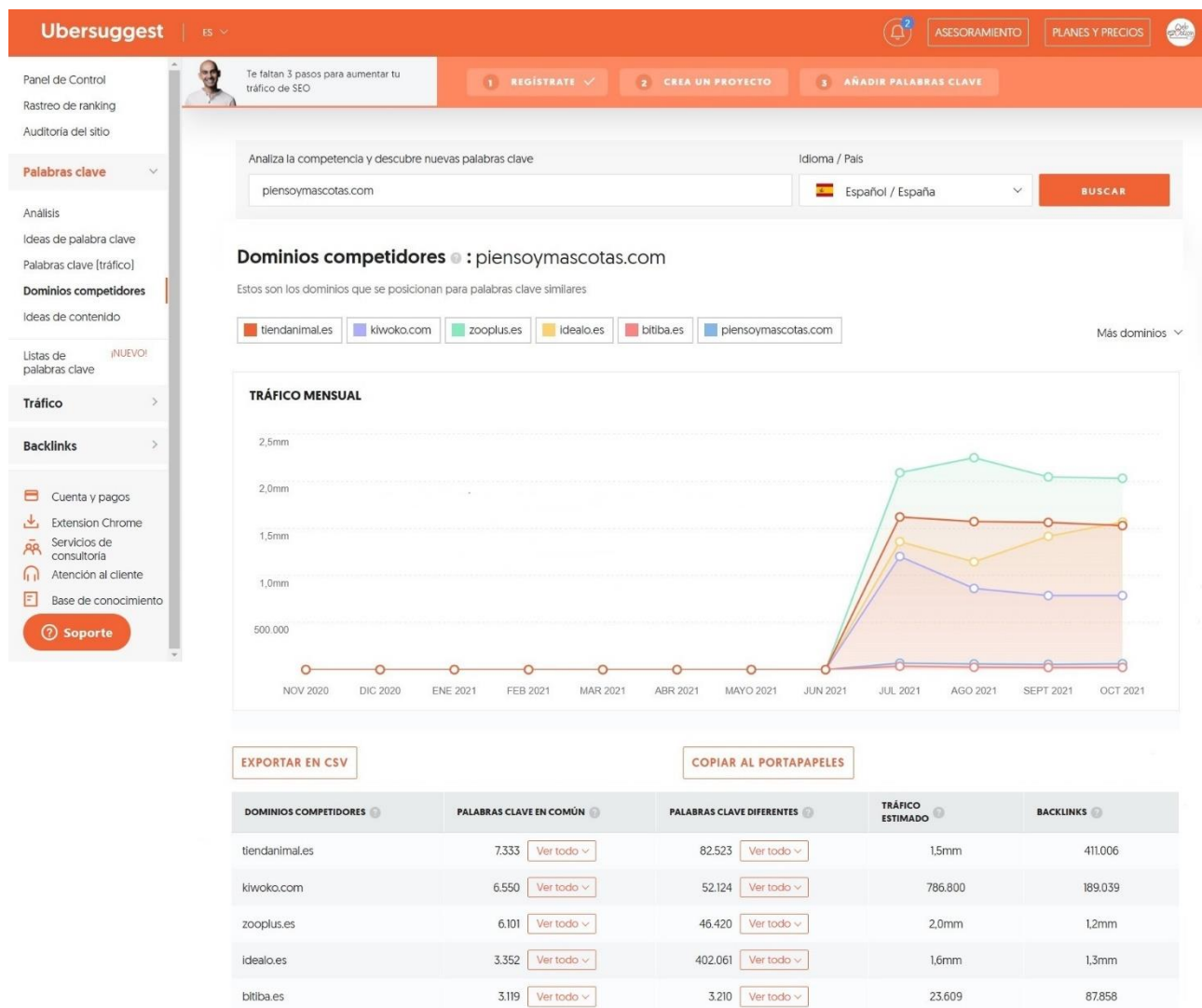


Ilustración 12. Vista de ejemplo de competidores del sitio web piensoy mascotas.com

Para cruzar todos estos datos y obtener un listado completo y uniforme de todos los dominios que compiten en cuanto a intenciones de búsqueda entre sí se utilizó *Python* y el entorno de trabajo proporcionado por *Google Colab*.

Python es un lenguaje de programación muy sencillo que se asemeja al lenguaje natural y que nos permite manejar grandes cantidades de datos; además empieza a ser ligeramente utilizado para automatizar tareas relacionadas con el análisis de sitios web.

En primer lugar, se hizo una preparación del entorno importando las librerías de *Python* necesarias para realizar los análisis y se accedió a las carpetas donde teníamos la información extraída de *Ubersuggest*.

```
import numpy as np
import pandas as pd
import os
import matplotlib.pyplot as plt

[ ] # Accedemos a la carpeta de google drive donde tenemos alojados los archivos
from google.colab import drive
drive.mount('/content/drive')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

[ ] # para ir a la carpeta donde tenemos los datos
os.chdir('/content/drive/MyDrive/Colab Notebooks/')
```

Ilustración 13. Preparación del entorno en *Python* y *Google Colab*.

A continuación, se cargaron todos los datos de los sitios web de tipología *ecommerce* seleccionados de la muestra sectorial.

```
[ ] #Leemos los datos del archivo csv donde hemos recopilado los principales competidores de cada uno de los ecommerce elegidos
# y los almacenamos en variables de tipo dataframe
dfCompetidoresBitiba = pd.read_csv('competidores/competitors_data_for_https__www.bitiba.es.csv', header = 'infer', encoding='utf8')
dfCompetidoresKiwoko = pd.read_csv('competidores/competitors_data_for_https__www.kiwoko.com.csv', header = 'infer', encoding='utf8')
dfCompetidoresMascotaPlanet = pd.read_csv('competidores/competitors_data_for_https__www.mascotaplanet.com.csv', header = 'infer', encoding='utf8')
dfCompetidoresMascoteros = pd.read_csv('competidores/competitors_data_for_https__www.mascoteros.com.csv', header = 'infer', encoding='utf8')
dfCompetidoresMiscota = pd.read_csv('competidores/competitors_data_for_https__www.miscota.es.csv', header = 'infer', encoding='utf8')
dfCompetidoresNuestroPerro = pd.read_csv('competidores/competitors_data_for_https__www.nuestroperro.es.csv', header = 'infer', encoding='utf8')
dfCompetidoresPetcllic = pd.read_csv('competidores/competitors_data_for_https__www.petcllic.es.csv', header = 'infer', encoding='utf8')
dfCompetidoresPiensoyMascotas = pd.read_csv('competidores/competitors_data_for_https__www.piensoymascotas.com.csv', header = 'infer', encoding='utf8')
dfCompetidoresTiendanimal = pd.read_csv('competidores/competitors_data_for_https__www.tiendanimal.es.csv', header = 'infer', encoding='utf8')
```

Ilustración 14. Carga de datos en *Python*.

Se creó una función en la que su finalidad era ir recopilando todos los competidores que tenía cada uno de los sitios web seleccionados y que fue almacenando junto con la cantidad de veces que aparecía cada competidor.

```
[ ] # Creamos una variable de tipo diccionario con los valores de los que hemos partido
# donde iremos almacenando los datos de frecuencia de competidores que aparecen en los
# datos que hemos cargado
diccionario = {
    "bitiba.es": 1,
    "kiwoko.com": 1,
    "mascotaplanet.com": 1,
    "mascoteros.com": 1,
    "miscota.es": 1,
    "nuestroperro.es": 1,
    "petclic.es": 1,
    "piensoymascotas.com": 1,
    "tiendanimal.es": 1
}

# creamos una función a la que le pasaremos cada dataframe
def creaDiccionario (df):
    # recorreremos por filas el dataframe
    for i in df.index:
        # para cada fila de la columna donde aparece la información del dominio de la competencia
        # si no está en el diccionario lo añadimos con valor uno por su primera aparición
        if df['Dominio de la competencia'][i] not in diccionario:
            diccionario[df['Dominio de la competencia'][i]] = 1
        # si ya está en el diccionario, entonces incrementamos en uno el valor
        else:
            diccionario[df['Dominio de la competencia'][i]] += 1

# a continuación pasamos cada uno de los dataframe por la función para que cree el diccionario
creaDiccionario(dfCompetidoresBitiba)
creaDiccionario(dfCompetidoresKiwoko)
```

Ilustración 15. Función que recopila todos los competidores.

De esta forma se obtuvo un listado mucho más amplio de sitios web que realmente compiten entre sí, hasta un total de 40 competidores y que se ordenaron por la cantidad de veces que se repetía su aparición.

```
[ ]
```

	Competidor	Apariciones
1	kiwoko.com	9
4	miscota.es	9
8	tiendanimal.es	9
9	zooplus.es	9
10	zoomalia.es	9
21	expertoanimal.com	7
12	elcorteingles.es	7

Ilustración 16. Muestra parcial de todos los competidores encontrados.

De nuevo se realizó una selección por tipología y se escogieron los *ecommerce* con la intención de obtener una muestra lo suficientemente uniforme para ser analizada.

Listado de <i>ecommerce</i>
animalear.com
barakaldotiendaveterinaria.es
bitiba.es
huellacanina.com
Kiwoko.com
mascotaplanet.com
mascoteros.com
maskokotas.com
miscota.es
nuestroperro.es
perrogato.net
petclic.es
petsfarma.es
piensosraposo.es
tiendanimal.es
zoomalia.es
zooplus.es

Comparando la muestra inicial de la partía con la muestra final se observa que algunos sitios fueron descartados y ocupados por otros que tras el análisis realizado arrojó que realmente eran competidores entre sí.

7.4 Análisis de palabras clave

Una vez elegida la muestra se llevó a cabo una investigación de las principales palabras clave por las que posicionan todos los sitios web. Para ello se utilizó la herramienta *Ubersuggest* de la que se extrajeron los datos de cada uno de los dominios en estudio para cruzarlos en *Python* y obtener las principales palabras clave en común por las que compiten.

Seoabilidad. Cómo aplicar el SEO centrado en la Usabilidad, Grado Multimedia, Patricia Luengo Carretero

Obtén una visión general del SEO de un sitio web o de una URL

Idioma / País: Español / España

https://www.bitiba.es

Análisis de tráfico : https://www.bitiba.es

- PALABRAS CLAVE ORGÁNICAS**: 14.447 BUENO
- TRÁFICO ORGÁNICO MENSUAL**: 23.609
- DOMAIN AUTHORITY**: 32 BUENO
- BACKLINKS**: 87.858 GENIAL (NoFollow: 1.061)

PALABRAS CLAVE DE SEO

PALABRAS CLAVE DE SEO	VOLUMEN	POSICIÓN	EST. VISITAS
bitiba gatos	880	1	404
bitiba	260	1	114
arenero para gatos	12.100	6	110
rascadores para gatos	18.100	6	109
cama gatos	4.400	5	102
areneros gatos	3.600	5	99
pienso perro	5.400	8	95
arenero gato	3.600	5	94
bitiba gmbh	210	1	94

Ilustración 17. Vista de ejemplo de obtención de palabras clave en el dominio bitiba.es

Se cargaron los listados de palabras clave para cada uno de los dominios analizados en Python y se guardó cada informe de palabras clave en una variable.

```
[ ] #Leemos los datos del archivo csv donde hemos recopilado las principales keywords por las que estan posicionadas las webs competidoras
dfKeywordsAnimalear = pd.read_csv('keywords/ubersuggest_www.animalear.com.csv', header = 'infer', encoding='utf8')
dfKeywordsBarakaldoTiendaVeterinaria = pd.read_csv('keywords/ubersuggest_www.barakaldotiendaveterinaria.es.csv', header = 'infer', encoding='utf8')
dfKeywordsBitiba = pd.read_csv('keywords/ubersuggest_www.bitiba.es.csv', header = 'infer', encoding='utf8')
dfKeywordsHuellaCanina = pd.read_csv('keywords/ubersuggest_www.huellacanina.com.csv', header = 'infer', encoding='utf8')
dfKeywordsKiwoko = pd.read_csv('keywords/ubersuggest_www.kiwoko.com.csv', header = 'infer', encoding='utf8')
dfKeywordsMascotaPlanet = pd.read_csv('keywords/ubersuggest_www.mascotaplanet.com.csv', header = 'infer', encoding='utf8')
dfKeywordsMascoteros = pd.read_csv('keywords/ubersuggest_www.mascoteros.com.csv', header = 'infer', encoding='utf8')
dfKeywordsMaskokotas = pd.read_csv('keywords/ubersuggest_www.maskokotas.com.csv', header = 'infer', encoding='utf8')
dfKeywordsMiscota = pd.read_csv('keywords/ubersuggest_www.miscota.es.csv', header = 'infer', encoding='utf8')
dfKeywordsNuestroPerro = pd.read_csv('keywords/ubersuggest_www.nuestroperro.es.csv', header = 'infer', encoding='utf8')
dfKeywordsPerroGato = pd.read_csv('keywords/ubersuggest_www.perrogato.net.csv', header = 'infer', encoding='utf8')
dfKeywordsPetcllic = pd.read_csv('keywords/ubersuggest_www.petcllic.es.csv', header = 'infer', encoding='utf8')
dfKeywordsPetsfarma = pd.read_csv('keywords/ubersuggest_www.petsfarma.es.csv', header = 'infer', encoding='utf8')
dfKeywordsPiensosRaposo = pd.read_csv('keywords/ubersuggest_www.piensosraposo.es.csv', header = 'infer', encoding='utf8')
dfKeywordsPiensosMascotas = pd.read_csv('keywords/ubersuggest_www.piensomascotas.com.csv', header = 'infer', encoding='utf8')
dfKeywordsTiendanimal = pd.read_csv('keywords/ubersuggest_www.tiendanimal.es.csv', header = 'infer', encoding='utf8')
dfKeywordsZoomalia = pd.read_csv('keywords/ubersuggest_www.zoomalia.es.csv', header = 'infer', encoding='utf8')
dfKeywordsZooplus = pd.read_csv('keywords/ubersuggest_www.zooplus.es.csv', header = 'infer', encoding='utf8')

# guardamos en una lista todos los archivos de datos a modo de indice
listaDatos = [dfKeywordsAnimalear, dfKeywordsBarakaldoTiendaVeterinaria, dfKeywordsBitiba, dfKeywordsHuellaCanina, dfKeywordsKiwoko, dfKeywordsMascoteros, dfKeywordsMaskokotas, dfKeywordsMiscota, dfKeywordsNuestroPerro, dfKeywordsPerroGato, dfKeywordsPetcllic, dfKeywordsPetsfarma, dfKeywordsPiensosRaposo, dfKeywordsPiensosMascotas, dfKeywordsTiendanimal, dfKeywordsZoomalia, dfKeywordsZooplus]
listaNombres = ['animalear.com', 'barakaldotiendaveterinaria.es', 'bitiba.es', 'huellacanina.com', 'kiwoko.com', 'mascotaplanet.com', 'nuestroperro.es', 'perrogato.net', 'petcllic.es', 'petsfarma.es', 'piensosraposo.es', 'piensomascotas.com', 'tiendanimal.es', 'zoomalia.es', 'zooplus.es']
```

Ilustración 18. Carga de datos de palabras clave de la muestra de dominios competidores.

Se realizó una comprobación de uno de los ficheros al azar para comprobar que los datos se habían cargado correctamente.

```
[ ] #Comprobamos como se ven los datos en los dataframes
dfKeywordsAnimalear
```

	No	Palabras clave	Volumen	Posición	Est Visitas	SEO Difficulty
0	1	animalear app	90	1	29	68
1	2	animalear tienda	70	1	29	43
2	3	animalear tienda fisica	40	1	14	43
3	4	animalear tienda física	40	1	14	43
4	5	animalear telefono	170	2	19	32
...
212	213	tienda de gatos sevilla	30	98	0	53
213	214	comprar gatos mallorca	50	100	0	51
214	215	ropa de animales online	0	100	0	4
215	216	tienda animales cordoba	390	100	0	39
216	217	tienda de animales en almeria	390	101	0	50

217 rows x 6 columns

Ilustración 19. Comprobación de la carga de datos.

Y, por último, se implementó una función que cruza todos los datos de todos los archivos buscando las palabras clave en común y la cantidad de veces que aparecen.

```
[ ] # Creamos una variable de tipo diccionario donde iremos guardando todas las palabras clave que hay entre todos los competidores
# y además iremos guardando el número de veces que se repite para localizar todos los competidores que compiten por la misma palabra clave
keywords = {}

# creamos una función a la que le pasaremos cada dataframe de cada dominio
def creaKeywords (dfk, n):
    # recorremos por filas el dataframe
    for i in dfk.index:
        # para cada fila de la columna donde aparece la información de la palabra clave
        # si no está en el diccionario lo añadimos con valor uno por su primera aparición y almacenamos también el volumen de búsquedas
        if dfk['Palabras clave'][i] not in keywords:
            keywords[dfk['Palabras clave'][i]] = {'Sitio Web': listaNombres[n], 'Aparición': 0, 'Volumen': dfk['Volumen'][i], 'Estimación Visitas':
            # si ya está en el diccionario, entonces incrementamos en uno el valor de la aparición
            # el volumen de búsquedas se deja como está pues va a ser el mismo para la misma keyword
        else:
            keywords[dfk['Palabras clave'][i]]['Aparición'] += 1

# a continuación pasamos cada uno de los dataframe por la función creada creaKeywords junto con el indice de webs para que cree el diccion
indice = 0
for lista in listaDatos:
    creaKeywords(lista, indice)
    indice += 1
```

Ilustración 20. Función que busca las palabras clave en común entre todos los sitios web.

Se visualiza el resultado y se devuelve ordenado en función de la cantidad de veces que aparecieron las palabras clave:

	Palabras Clave	Aparición	Volumen
388	comprar pienso perros online	16	70
31	tienda para perros	16	720
168	tienda online para perros	16	90
23	tienda perros	16	1900
468	pienso perros online	16	90
423	comprar pienso perro online	16	70
42	tiendas de perros	16	2400
46	tienda online de mascotas	15	480
56	tienda animales gatos	15	70
833	tiendas mascotas cerca de mi	15	720

Se observó que varias palabras clave se repetían entre los 16 de los 18 competidores que se analizaron. Entre ellas la palabra clave “tiendas de perros” que es la que más volumen de búsquedas mensuales tiene, así que es la que se eligió para buscar el ranking que cada sitio web tiene para esa intención de búsqueda.

De nuevo se creó una función que recorría todos los archivos de datos buscando la palabra clave escogida y obteniendo su ranking.

```

▶ # Creamos una variable de tipo diccionario donde iremos almacenando todos los datos extraidos
ranking = {}
# Guardamos en una variable nuestra palabra clave objetivo
keyword = 'tiendas de perros'
# creamos una función a la que le pasaremos cada archivo de datos de cada uno de los competidores
def creaRanking (df, n):
    # recorremos por filas el dataframe
    for i in df.index:
        if df['Palabras clave'][i] == keyword:
            ranking[df['Posición'][i]] = {'Sitio Web': listaNombres[n], 'Palabras Clave': df['Palabras clave'][i].

indice = 0
for lista in listaDatos:
    creaRanking(lista, indice)
    indice += 1

ranking
    
```

Ilustración 21. Obtención del ranking para la búsqueda "tiendas de perros"

Se muestra el resultado por orden de posición en el ranking de búsquedas, visualizando los diez primeros resultados que se corresponden con los sitios web a auditar:

	Sitio Web	Palabras Clave	Posición
15	zooplus.es	tiendas de perros	1
4	kiwoko.com	tiendas de perros	2
7	miscota.es	tiendas de perros	3
13	tiendanimal.es	tiendas de perros	4
6	maskokotas.com	tiendas de perros	6
5	mascoteros.com	tiendas de perros	8
14	zoomalia.es	tiendas de perros	9
2	bitiba.es	tiendas de perros	14
10	petclic.es	tiendas de perros	15
8	nuestroperro.es	tiendas de perros	21

Ilustración 22. Ranking de los diez sitios web a auditar

7.5 Métricas de la muestra elegida

En este apartado se realizó un análisis básico de los dominios escogidos en base al marco teórico expuesto en el que se han tomado las métricas más relevantes que hacen referencia al SEO y la usabilidad y que pueden ser extraídas de forma externa a los propietarios de los sitios web.

Indexabilidad

Se recopilaron datos de **url amigables** que son aquellas que resultan entendibles a los usuarios y al buscador en su tendencia a dirigirse a una web semántica.

Los datos se extrajeron con *Ubersuggest* y suponen una métrica en porcentaje. La herramienta muestra cuantas páginas no tienen url amigables y también nos da

información de cuantas páginas ha encontrado en el sitio, de esta forma podemos obtener el porcentaje de cuantas url amigables hay en el sitio web y en la que se obtuvieron los siguientes resultados.

Listado de <i>ecommerce</i>	Url amigables %
bitiba.es	43,11
kiwoko.com	73,80
mascoteros.com	99,70
maskokotas.com	100,00
miscota.es	100,00
nuestroperro.es	94,70
petclic.es	100,00
tiendanimal.es	99,90
zoomalia.es	98,74
zooplus.es	13,80

Tabla 3. Porcentaje de url amigables

Los datos de ***thin content*** constituyen una métrica que no solo afecta a la indexabilidad del sitio web, sino que también tiene que ver con la calidad del contenido. Si una página del sitio web no tiene calidad suficiente y no aporta nada al usuario, el buscador la considera pobre y no la indexa, por tanto, no aparecerá en los resultados de búsqueda.

Los datos se obtuvieron de *Ubersuggest* y se presentan en porcentaje tomando la cantidad de páginas consideradas de baja calidad frente a la cantidad de páginas rastreadas y en la que se obtuvieron los siguientes resultados.

Listado de <i>ecommerce</i>	<i>Thin content</i> %
bitiba.es	7,32
kiwoko.com	0,01
mascoteros.com	0,01
maskokotas.com	0,05
miscota.es	0,06
nuestroperro.es	0,01
petclic.es	0,01
tiendanimal.es	0,01
zoomalia.es	0,06

zooplus.es	14,80
------------	-------

Tabla 4. Porcentaje de páginas con *thin content*.

La existencia de **enlaces rotos** no sólo supone una falta de usabilidad para el usuario que puede frustrarse cuando se encuentra con una página que no existe, sino que son páginas que si las detecta el buscador son desindexadas.

Los datos obtenidos a través de *Ubersuggest* y expresados en porcentaje en función de la cantidad de páginas rastreadas por dominio fueron los siguientes:

Listado de ecommerce	Enlaces rotos %
bitiba.es	0,01
kiwoko.com	0,01
mascoteros.com	0,01
maskokotas.com	0,54
miscota.es	0,87
nuestroperro.es	0,01
petclic.es	1,61
tiendanimal.es	0,01
zoomalia.es	1,91
zooplus.es	2,30

Tabla 5. Porcentaje de enlaces

Contenido

La presencia de **etiquetas** correctamente etiquetadas en los atributos de imágenes ALT, una buena jerarquía en las etiquetas de tipo H1 a H6, los *title* de las páginas y sus *meta descriptions* no sólo facilitan la comprensión al usuario y mejoran su accesibilidad, sino que son bien valorados por los buscadores.

De esta forma se tomó como métrica la cantidad de páginas en porcentaje que están correctamente etiquetadas mediante la información que proporciona *Ubersuggest* y que dio como resultado las siguientes medidas:

Listado de <i>ecommerce</i>	Etiquetas Ok %
bitiba.es	49,21
kiwoko.com	94,60
mascoteros.com	71,20
maskokotas.com	65,20
miscota.es	75,30
nuestroperro.es	86,40
petcllic.es	5,54
tiendanimal.es	87,21
zoomalia.es	84,09
zooplus.es	46,20

Tabla 6. Porcentaje de páginas correctamente etiquetadas

El **CTR (Click Through Rate)** es la proporción de clics que recibe un enlace con respecto a la cantidad de veces que ha aparecido en las búsquedas.

Para hacer el cálculo real se necesitaría tener acceso a los sitios web y consultarlos con *Google Search Console*, así que se obtuvo una estimación con los datos obtenidos en *UberSuggest* y *Python*. Para la palabra clave elegida se tomó la estimación de visitas para cada uno de los casos de estudio entre la estimación del volumen de apariciones en el buscador expresado en porcentaje con los siguientes resultados:

[]	Sitio Web	URL destino	Enlaces	Autoridad de Dominio	Visitas	CTR
0	bitiba.es	http://www.bitiba.es/	86.582	32	3	0.12
1	kiwoko.com	http://www.kiwoko.com/	18.618	54	729	30.38
2	mascoteros.com	http://www.mascoteros.com/	5.694	48	0	0.00
3	maskokotas.com	http://www.maskokotas.com/	267.000	35	47	1.96
4	miscota.es	http://www.miscota.es/	50.959	43	233	9.71
5	nuestroperro.es	http://www.nuestroperro.es/	6.385	32	2	0.08
6	petcllic.es	http://www.petcllic.es/	35.264	36	0	0.00
7	tiendanimal.es	http://www.tiendanimal.es/	123.076	64	158	6.58
8	zoomalia.es	http://www.zoomalia.es/	200.423	30	61	2.54
9	zooplus.es	http://www.zooplus.es/	297.657	52	388	16.17

Ilustración 23. Datos de estimación de CTR

La **autoridad del dominio** es una métrica externa creada por la empresa *Moz* y que se puede obtener mediante una extensión del navegador *Chrome*. Esta métrica se basa en un puntaje de 1 a 100 y proporciona información acerca de la autoridad que un dominio puede tener.

Listado de <i>ecommerce</i>	Autoridad del Dominio DA
bitiba.es	32
kiwoko.com	54
mascoteros.com	48
maskokotas.com	35
miscota.es	43
nuestroperro.es	32
petclic.es	36
tiendanimal.es	64
zoomalia.es	30
zooplus.es	52

Tabla 7. Autoridad de dominio.

Las **páginas por visita** es una métrica con la que se valora si el contenido que ofrece el sitio web resulta de utilidad para el usuario o no, cuantas más páginas visite en el sitio mejor. Los datos que se muestran a continuación fueron obtenidos utilizando la herramienta de análisis de tráfico *Similarweb*:

Listado de <i>ecommerce</i>	Páginas por visita
bitiba.es	6,26
kiwoko.com	4,44
mascoteros.com	2,1
maskokotas.com	3,05
miscota.es	4,05
nuestroperro.es	1,98
petclic.es	3,49
tiendanimal.es	3,97
zoomalia.es	4,62
zooplus.es	5,56

Tabla 8. Páginas por visita

La **duración media de la visita o tiempo de permanencia en el sitio** es también un signo de calidad del sitio web. Las métricas para los sitios web analizados se obtuvieron con *Similarweb* y se muestran a continuación:

Listado de <i>ecommerce</i>	Duración media de la visita (s)
bitiba.es	301
kiwoko.com	172
mascoteros.com	64
maskokotas.com	99
miscota.es	165
nuestroperro.es	241
petclic.es	102
tiendanimal.es	154
zoomalia.es	246
zooplus.es	276

Tabla 9. Duración media de la visita

Por el contrario, la **tasa de rebote** es una métrica que proporciona información acerca del porcentaje de personas que entraron el sitio web pero que no pasaron de la primera página. Si es un porcentaje muy alto puede indicar que el contenido del sitio web no es bueno.

Listado de <i>ecommerce</i>	Tasa de rebote %
bitiba.es	48,07
kiwoko.com	67,48
mascoteros.com	71,33
maskokotas.com	63,05
miscota.es	56,25
nuestroperro.es	73,87
petclic.es	55,78
tiendanimal.es	65,29
zoomalia.es	51,83
zooplus.es	48,94

Tabla 10. Tasa de rebote

Mobile First

Desde 2020, *Google* construye su índice a partir del rastreo realizado dando prioridad a la versión móvil así que se introdujo en el análisis la velocidad de las páginas en estos dispositivos a través de la aplicación online, *Test my site* que *Google* pone a disposición de los desarrolladores.

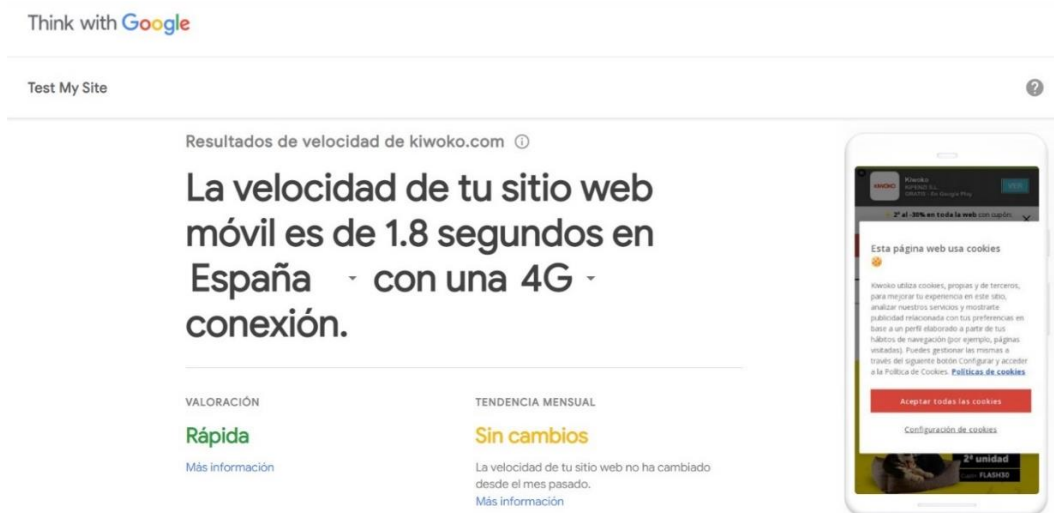


Ilustración 24. Ejemplo de captura de datos de velocidad móvil con *Test my site*

Los resultados obtenidos para todos los sitios analizados fueron:

Listado de <i>ecommerce</i>	Velocidad (s)
bitiba.es	2,50
kiwoko.com	1,80
mascoteros.com	2,40
maskokotas.com	2,5
miscota.es	2,7
nuestroperro.es	1,90
petclic.es	3,50
tiendanimal.es	2,40
zoomalia.es	2,90
zooplus.es	2,20

Tabla 11. Velocidad móvil

WPO

La optimización **WPO (Web Performance Optimization)** es un conjunto de técnicas que tratan de que el sitio web mejore en cuanto a su velocidad de carga.

La optimización del rendimiento web no solo es un factor clave en el posicionamiento, hay que tener en cuenta que la mayor parte del tráfico proviene de dispositivos móviles con unos tamaños de pantalla reducidos y con un menor ancho de banda.

Es por ello por lo que, a través de la herramienta de *Google, PageSpeed Insights*, se facilitan diversas métricas relacionadas con el rendimiento y la usabilidad.

Aunque el *WPO* está relacionado exclusivamente con las técnicas para mejorar la velocidad de carga de un sitio web, en esta herramienta *Google* está valorando mucho más la usabilidad y la experiencia de usuario y no sólo los datos de carga desde que introdujo las métricas relacionadas con las **Core Web Vitals** tal y como apunta (Fontela, 2021)

PageSpeed Insights

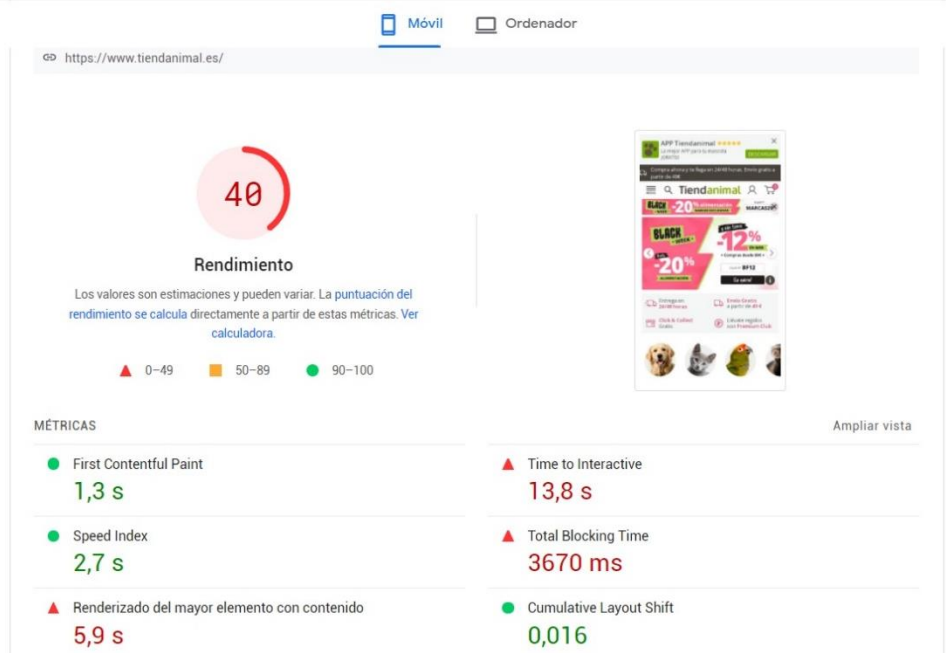


Ilustración 25. Ejemplo de captura de datos de WPO

Las métricas principales son:

- **Rendimiento:** Una métrica de rendimiento global del sitio web.
- **FCP (*First Contentful Paint*):** Mide el tiempo que tarda en cargar el contenido que se ve en pantalla sin necesidad de hacer scroll.
- **LCP (*Largest Contentful Paint*):** Mide el tiempo que tarda en cargar el elemento mayor que se ve en pantalla sin necesidad de hacer scroll.
- **Time to Interactive:** Mide el tiempo que tarda una web en ser totalmente interactiva para el usuario.
- **CLS (*Cumulative Layout Shift*):** Mide la estabilidad visual. Cuando aterrizamos en una página los elementos se van cargando, dado que esta carga no se produce a la vez, los elementos se van moviendo y a veces produce clics erróneos por parte del usuario. Así que lo que hace esta métrica es medir la frecuencia y magnitud con la que se producen estos cambios. La puntuación será mayor cuando más movimientos indeseados se registren.

Popularidad

Esta parte de la auditoría es la que dista completamente de la usabilidad pero que ha de tenerse en cuenta para valorar su peso en el ranking.

Se trata de analizar si el sitio web es relevante y se mide en la cantidad de enlaces que el dominio a estudiar recibe por parte de otros dominios; se basa en la idea de la referencia, ya que si otros te nombran es que eres referente.

Los **enlaces totales** se extrajeron con *Ubersuggest* y es la medida de la cantidad de enlaces que apuntan al dominio sin tener en cuenta si todos proceden de un solo dominio o de varios.

Listado de <i>ecommerce</i>	Enlaces totales
bitiba.es	151
kiwoko.com	798
mascoterros.com	1.715
maskokotas.com	723
miscota.es	796
nuestroperro.es	216
petcllc.es	229
tiendanlmal.es	1.734
zoomalia.es	209
zooplus.es	587

Tabla 12. Enlaces totales

Los **enlaces *follow*** son aquellos enlaces que han sido etiquetados así por parte del dominio que apunta al del sitio analizado; son importantes porque traspasan autoridad.

Listado de <i>ecommerce</i>	Enlaces <i>follow</i>
bitiba.es	12
kiwoko.com	142
mascoterros.com	1.534
maskokotas.com	404
miscota.es	51
nuestroperro.es	29
petcllc.es	43
tiendanlmal.es	247
zoomalia.es	1
zooplus.es	74

Tabla 13. Enlaces follow

Los **enlaces no *follow*** son similares a los anteriores sin embargo estos no traspasan autoridad.

Listado de <i>ecommerce</i>	Enlaces no <i>follow</i>
bitiba.es	139
kiwoko.com	656
mascoteros.com	181
maskokotas.com	319
miscota.es	745
nuestroperro.es	187
petclic.es	186
tiendanimal.es	1.487
zoomalia.es	208
zooplus.es	513

Tabla 14. Enlaces no follow

Los **dominios únicos** de referencia suponen la cantidad de dominios únicos que tienen enlaces apuntando al sitio web analizado.

Listado de <i>ecommerce</i>	Dominios únicos
bitiba.es	43
kiwoko.com	152
mascoteros.com	101
maskokotas.com	133
miscota.es	74
nuestroperro.es	49
petclic.es	68
tiendanimal.es	234
zoomalia.es	16
zooplus.es	159

Tabla 15. Dominios únicos de referencia

7.6 Modelo empleado y resultados del análisis

Se utilizó *Python* para analizar los datos obtenidos y después se aplicó un modelo de regresión lineal múltiple para tratar de dar cierta veracidad a todos los datos analizados.

En primer lugar, se centralizaron todos los datos recabados para el estudio en una hoja Excel que se cargó en Python.

```
# Accedemos a la carpeta de google drive donde tenemos alojados los archivos
from google.colab import drive
drive.mount('/content/drive')

Drive already mounted at /content/drive; to attempt to forcibly remount, call drive.mount("/content/drive", force_remount=True).

# para ir a la carpeta donde tenemos los datos
os.chdir('/content/drive/MyDrive/Colab Notebooks/')

#Leemos los datos del archivo excel donde hemos recopilado todos las métricas de los 10 sitios elegidos y lo almacenamos en una variable de t
dfDatos = pd.read_excel('/content/drive/MyDrive/Colab Notebooks/Datos.xlsx', header = 'infer', sheet_name='Hoja2')
dfDatos
```

	Listado de ecommerce	Ranking	Url_amigables	Thin_content	Enlaces_rotos	Etiquetas_OK	CTR	DA	Velocidad	Rendimiento	FCP	LCP	CLS	TTI
0	bitiba.es	14	43.109756	7.317073	0.010000	49.207317	0.12	32	2.5	23	2.8	10.5	0.114	12.4
1	kiwoko.com	2	73.800000	0.010000	0.010000	94.600000	30.38	54	1.8	25	3.8	12.2	0.557	13.3
2	mascoteros.com	8	99.702085	0.010000	0.010000	71.201589	0.01	48	2.4	38	1.9	2.7	0.030	11.6
3	maskokotas.com	6	100.000000	0.050000	0.540000	65.200000	1.96	35	2.5	24	3.5	6.6	0.296	13.6
4	miscota.es	3	100.000000	0.060000	0.870000	75.300000	9.71	43	2.7	31	4.0	4.4	0.134	18.1
5	nuestroperro.es	21	94.700000	0.010000	0.010000	86.400000	0.08	32	1.9	45	3.9	6.8	0.481	6.6
6	petclie.es	15	100.000000	0.010000	1.608579	5.540661	0.01	36	3.5	52	2.8	2.8	0.101	10.5
7	tiendanimal.es	4	99.901961	0.010000	0.010000	87.205882	6.58	64	2.4	40	1.3	5.9	0.016	13.8

Ilustración 26. Vista parcial de la carga de datos en Python.

Gracias a la biblioteca *SciPy* utilizada en *Python* se pudo calcular fácilmente la matriz de correlación de *Pearson* que proporcionó información acerca de la relación que tienen entre si todas las métricas recogidas y que se transformaron en un mapa de calor para facilitar su interpretación.

Seoabilidad. Cómo aplicar el SEO centrado en la Usabilidad, Grado Multimedia, Patricia Luengo Carretero

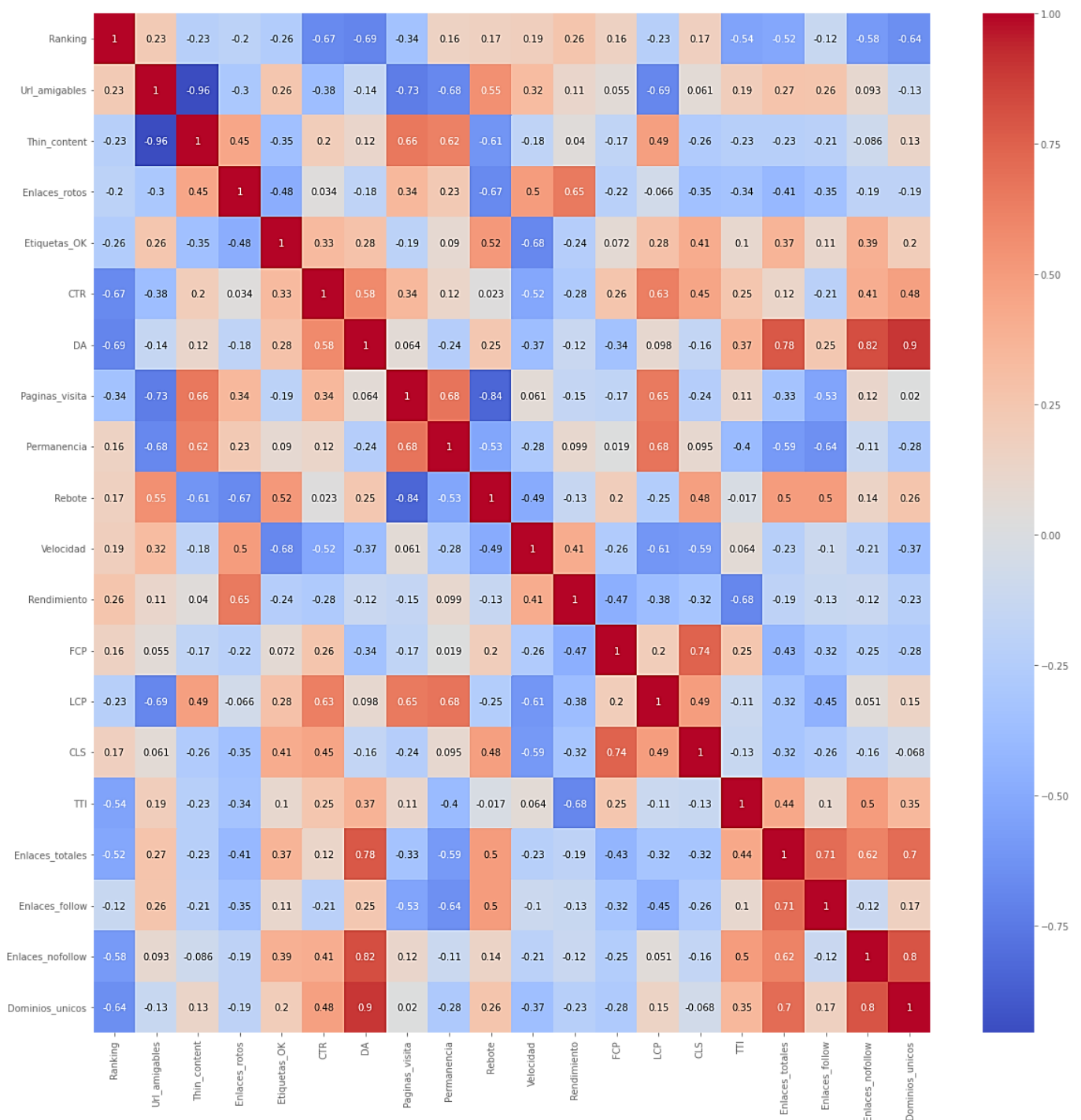


Ilustración 27. Matriz de correlación.

Si nos fijamos en la primera fila en la que se sitúa el ranking y vamos analizando columna a columna podemos interpretar como le afectan el resto de las métricas.

En primer lugar, destacan como correlaciones fuertes el CTR, la autoridad del dominio, el TTI (*Time to interactive*), los enlaces totales, los enlaces no *follow* y la cantidad de

dominios únicos. Son correlaciones negativas, es decir, que cuando el ranking aumenta, en valor numérico, el resto de las métricas disminuyen, o al revés.

A medida que disminuye el ranking en valor numérico, es decir tenemos mejores posiciones, el CTR aumenta en valor numérico, así que tenemos un CTR mejor; la autoridad de dominio también se incrementa en valor numérico, así que tenemos una mayor autoridad; con los enlaces sucede lo mismo, tenemos una mayor cantidad de enlaces cuanto mejor es el ranking.

Por otro lado, a medida que aumenta el ranking en valor numérico, es decir tenemos peores posiciones, tenemos una correlación positiva con la mayor parte de métricas relacionadas con la velocidad, es decir, tenemos velocidades más altas de carga en la web.

Solamente hay dos métricas que no siguen la lógica esperada, que son las url amigables y el TTI, ya que, según esta matriz de correlación a mejor posición peores datos.

Para validar en parte estos resultados se ha implementado un modelo matemático que se basa en el análisis de regresión lineal múltiple. Es un método estadístico que trata de relacionar una variable dependiente frente a múltiples variables independientes mediante el ajuste de una ecuación lineal.

Modelo matemático

Dado un conjunto de observaciones la variable respuesta o dependiente y se relacionará de forma lineal con las variables predictoras o independientes $x_1 \dots x_n$ en base a la siguiente ecuación:

$$y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon_i$$

Donde β_0 es la media de la variable respuesta y cuando todas las variables predictoras son cero y es la ordenada en el origen de la ecuación de regresión lineal múltiple

β_j son los llamados coeficientes parciales de regresión y se interpretan como el efecto que tiene sobre la variable respuesta el incremento en una unidad de su variable predictora correspondiente manteniéndose constantes el resto.

ε_i se denomina residuo y es la diferencia entre el valor observado y el estimado.

Implementación en Python

En nuestro caso la variable a modelar o variable respuesta será el ranking y el resto de las variables serán las variables independientes o predictores que, en este caso, una vez analizada la matriz de correlación vemos que se corresponden con el CTR, la autoridad del dominio, los enlaces totales, los enlaces no *follow* y la cantidad de dominios únicos de referencia.

Así pues, nos quedamos por un lado con las variables predictoras

```
[52] # Visualizamos los datos de X para comprobar que se ha realizado X
```

	CTR	DA	Enlaces_totales	Enlaces_nofollow	Dominios_unicos
0	0.12	32	151	139	43
1	30.38	54	798	656	152
2	0.01	48	1715	181	101
3	1.96	35	723	319	133
4	9.71	43	796	745	74
5	0.08	32	216	187	49
6	0.01	36	229	186	68
7	6.58	64	1734	1487	234
8	2.54	30	209	208	16
9	16.17	52	587	513	159

Ilustración 28. Variables predictoras.

Para utilizar correctamente este modelo se normalizaron los datos de las variables predictoras:

```
#Normalizamos los datos de las variables predictoras
scaler = MinMaxScaler()
scaled_values = scaler.fit_transform(X)
X.loc[:, :] = scaled_values

# Y visualizamos
X
```

	CTR	DA	Enlaces_totales	Enlaces_nofollow	Dominios_unicos
0	0.003622	0.058824	0.000000	0.000000	0.123853
1	1.000000	0.705882	0.408718	0.383531	0.623853
2	0.000000	0.529412	0.987997	0.031157	0.389908
3	0.064208	0.147059	0.361339	0.133531	0.536697
4	0.319394	0.382353	0.407454	0.449555	0.266055
5	0.002305	0.058824	0.041061	0.035608	0.151376
6	0.000000	0.176471	0.049274	0.034866	0.238532
7	0.216332	1.000000	1.000000	1.000000	1.000000
8	0.083306	0.000000	0.036639	0.051187	0.000000
9	0.532104	0.647059	0.275426	0.277448	0.655963

Ilustración 29. Normalización variables predictoras.

Y por otro lado se visualiza la variable respuesta o ranking.

```

0    14
1     2
2     8
3     6
4     3
5    21
6    15
7     4
8     9
9     1
Name: Ranking, dtype: int64

```

Ilustración 30. Variable respuesta.

Gracias a la biblioteca *scikit-learn* se pudieron realizar todos los cálculos necesarios para el modelo de regresión lineal múltiple en una sola línea código.

Para ello, antes se realizó un paso más, ya que es necesario dividir todo nuestro conjunto de datos, es decir, todas las observaciones para los 10 sitios web, en dos partes. Una parte, en concreto el 80 por ciento, es decir, los datos tomados para 8 sitios webs escogidos al azar, se utilizaron para entrenar el modelo, y el resto se utilizaron para test.

A continuación, se muestra todo este proceso llevado a cabo en Python.

```

[67] # Creamos dos conjuntos, uno para entrenamiento (80%) y otro para test(20%)
      X_train, X_test, Y_train, Y_test = train_test_split(X, Y.values.reshape(-1,1), test_size= 0.2, random_state = 123)

[68] # Visualizamos que se ha realizado correctamente
      print(X_train.shape)
      print(X_test.shape)
      print(Y_train.shape)
      print(Y_test.shape)

      (8, 5)
      (2, 5)
      (8, 1)
      (2, 1)

[69] # Cargamos el modelo
      modelo = LinearRegression()
      #Entrenamos
      modelo.fit(X_train,Y_train)
      # Predecimos en el entrenamiento
      predY_in = modelo.predict(X_train)
      # Predecimos en el test
      predY_out = modelo.predict(X_test)

```

Ilustración 31. Implementación del modelo.

El siguiente paso es obtener el *intercept* u ordenada en el origen de la ecuación de la recta que buscamos, así como los coeficientes de regresión lineal múltiple.

```
# Intercept
modelo.intercept_

array([15.86448846])

# Coeficientes, reflejan el impacto de cada una de las variables sobre el ranking
modelo.coef_

array([[ -12.00944601,   7.73828103,  -7.22583921,   2.67968799,
        -12.72503449]])

# Calculamos el coeficiente de determinación (R^2):
modelo.score(X,Y)

0.6059701099208674
```

Ilustración 32. Obtención de intercept, coeficientes de regresión y valor de R.

El modelo con todos los predictores introducidos tiene un R^2 de 0.6059, es decir, es capaz de explicar el 60.59% de la variabilidad observada en el ranking.

Y la ecuación de la recta que buscamos sería:

$$\text{Ranking} = 15,86 - 12,01 \text{ CTR} + 7.74 \text{ DA} - 7.23 \text{ EnlacesTotales} + 2.68 \text{ EnlacesNoFollow} - 12.73 \text{ DominiosUnicos}$$

La magnitud de los coeficientes parciales depende solo de las unidades en las que se han medido cada una de ellas así que estos coeficientes no nos están indicando la importancia de cada uno de ellos.

Por ejemplo, el decremento en unidad en el CTR, es decir, el CTR empeora, hace que el ranking incremente en 12, es decir, perdemos posiciones, manteniéndose constantes el resto de variables predictoras.

Ilustración 33. Datos estadísticos de las variables predictoras.

Finalmente, queríamos saber si el modelo creado a partir de todo este análisis era capaz de predecir la posición de una web en función de los datos de CTR, autoridad del dominio, enlaces totales, enlaces no *follow* y cantidad de dominios únicos de referencia.

Para ello, se utilizaron los datos de otro ecommerce lucasyloa.es que posiciona para la palabra clave “tiendas de perros” y que tiene las siguientes métricas:

- **CTR:** 3.375
- **DA:** 26
- **Enlaces totales:** 176
- **Enlaces no *follow*:** 71
- **Dominios únicos:** 59
- **Ranking:** 10

Antes de introducirlas en el modelo, estas métricas han de ser normalizadas, para ello se utilizan los datos estadísticos de las variables predictoras.

```
[8] # Obtenemos los datos estadísticos
dfDatos.describe(include='all')
```

	Ranking	Url_amigables	Thin_content	Enlaces_rotos	Etiquetas_OK	CTR	DA
count	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000	10.000000
mean	8.300000	82.375782	2.233688	0.728245	66.494636	6.756000	42.600000
std	6.532823	30.326034	4.974335	0.899755	26.775188	9.865049	11.500725
min	1.000000	13.800000	0.010000	0.010000	5.540661	0.010000	30.000000
25%	3.250000	79.025000	0.010000	0.010000	53.205488	0.090000	32.750000
50%	7.000000	99.223052	0.030000	0.275000	73.250794	2.250000	39.500000
75%	12.750000	99.975490	0.059952	1.423934	85.822727	8.927500	51.000000
max	21.000000	100.000000	14.800000	2.300000	94.600000	30.380000	64.000000

Ilustración 34. Vista parcial de datos estadísticos.

En concreto los datos de máximos y mínimos y se aplica la normalización utilizando la siguiente formula:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

A modo de ejemplo se normaliza la variable predictora CTR del dominio cuyos datos se introdujeron en el modelo para realizar la predicción del ranking

$$X_{CTR} = \frac{X - X_{min}}{X_{max} - X_{min}} = \frac{3.375 - 0.01}{30.38 - 0.01} = 0.1039$$

Quedando los datos normalizados de la siguiente forma:

- **CTR:** 0.1039
- **DA:** 0
- **Enlaces totales:** 0.015
- **Enlaces no *follow*:** 0
- **Dominios únicos:** 0.1972
- **Ranking:** 10

Que al introducir en el modelo dieron como resultado el siguiente puesto de ranking

```
[ ] # Tenemos que normalizar los datos antes de introducirlos en el modelo para hacer la predicción
X_nuevo = np.array([0.1039, 0, 0.015, 0, 0.1972]).reshape(1,-1)
X_nuevo

array([[0.1039, 0.    , 0.015 , 0.    , 0.1972]])
```

```
[ ] Y_pred_nuevo = modelo.predict(X_nuevo)
print('La predicción de ranking para lucasyloa.es es:', np.round(Y_pred_nuevo[0,0],3))

La predicción de ranking para lucasyloa.es es: 11.999
```

Ilustración 35. Predicción.

El modelo ha predicho que el puesto que ocupará con esos datos es el 12, vemos que aún a pesar de la poca cantidad de observaciones que tenemos para emplear estos modelos ha hecho una buena predicción lo que nos ayuda a dar veracidad a todo el análisis realizado.

Observaciones en cuanto al modelo

Son necesarias muchas más observaciones. Dependiendo del número de predictores que vayamos a utilizar serían necesarias de diez a veinte observaciones más por cada predictor. En nuestro modelo se han usado cinco predictores así que para dar los resultados como válidos habríamos necesitado hasta cincuenta observaciones.

En los modelos de regresión lineal múltiple, las variables predictoras han de ser independientes entre sí y aunque no hay un método matemático que revele la existencia de colinealidad entre las variables, la matriz de correlación nos da una idea de su existencia, aunque aun así es difícil de afirmar. Por ejemplo, si analizamos la matriz de correlación vemos que el número de dominios únicos está fuertemente correlacionado con el dominio de autoridad, así como el contenido *thin content* y las URL amigables o la tasa de rebote con las páginas vistas.

Otro punto para considerar son los datos, la realidad es que en la mayor parte de los casos son estimaciones por parte de herramientas de terceros y que lo ideal sería tener acceso a los datos reales que pertenecen a los propietarios de los sitios web.

8. Metodología para Seobilidad

En 2012 Jakob Nielsen, referente en cuanto a la usabilidad web, describía en su artículo *SEO and Usability*, la importancia de la relación entre usabilidad y SEO afirmando que “lo que hace que un sitio web sea bueno también le dará un alto rango de SERP, pero una optimización de motor de búsqueda demasiado complicada puede socavar la experiencia del usuario.” (Nielsen, Seo and usability, 2012). También podemos encontrar algunas investigaciones en torno al SEO y la usabilidad como el realizado para evaluar estos factores en ochenta sitios web de universidades árabes. (Al-Ananbeh, 2012)

Por otro lado, Fernando Maciá, referente en cuanto a posicionamiento de sitios web introduce en su libro “SEO Avanzado, casi todo lo que sé sobre posicionamiento web” diversas referencias al diseño pensando en el usuario y no sólo en el motor de búsqueda.

Finalmente, el propio motor de búsqueda introdujo en 2020 métricas relacionadas con la experiencia de usuario. Según Bryan McQuade, en su artículo “Definición de los umbrales de las métricas de Core Web Vitals”, estas tienen como objetivo “garantizar que las páginas cumplen los umbrales "buenos" de Core Web Vitals y ofrezcan una experiencia de usuario de alta calidad.” (McQuade, Definición de los umbrales de las métricas de Core Web Vitals, 2020)

Todo ello hace que sea razonable plantear ambas disciplinas como una sola y es a partir del marco teórico y práctico desarrollado que se propone a continuación una metodología a seguir.

8.1. Análisis

En esta fase se establecen las bases del sitio web en cuanto al propósito, el estudio de la competencia, los requerimientos técnicos y del cliente y el público objetivo al que vamos a dirigirnos.

- **Definición.** Se especificará el propósito del sitio web, la tipología, temática principal y secundarias.
- **Objetivos.** Estarán directamente relacionados con la definición del sitio web y su tipología, por ejemplo, vender productos, conseguir suscriptores, crear imagen de marca, generar leads, crear comunidad, etc.
- **Requerimientos del cliente.** Es necesaria su participación en esta fase, ya que nos dará una idea inicial de cuál es su negocio, a quien va dirigido, cual es el estilo, qué otras páginas le gustan o no, cuál es su competencia, etc.
- **Requerimientos técnicos.** Relativo a las tecnologías que se usarán para construir el sitio, desde un desarrollo a medida al uso de un CMS así como la elección del *hosting*.
- **Estudio de la competencia.** Utilizando una selección de los términos principales del sitio web podemos identificar a los dominios competidores y nos dará una idea del mercado al que vamos a dirigirnos, qué hace la competencia y cómo para incorporar su estrategia a la nuestra, mejorarla o que nos sirva de inspiración.
- **Características de los usuarios.** Conocer la audiencia o público objetivo al que nos dirigimos es fundamental para conocer las necesidades e intereses de los usuarios.

8.2. Modelado del usuario

8.2.1. Personas

El proceso de creación de personas nos ayudará a empatizar con el público objetivo del sitio web, identificar sus necesidades, comportamientos, experiencias, miedos, objeciones, etc. y se utiliza en el modelado del usuario en el DCU; si además incluimos los aspectos utilizados en el ámbito SEO podremos obtener un arquetipo mucho más enriquecido.

En una ficha de persona se deberán incluir los siguientes aspectos:

- **Imagen y nombre:** Cuando ponemos cara al usuario nos resulta más sencillo empatizar con él.
- **Frase** con la que se define.
- **Datos demográficos:** Edad, sexo, localización geográfica, nivel educativo, situación familiar, capacidad adquisitiva
- **Definición de la personalidad**
- **Biografía**
- **Trabajo y cargo:** Se describe el sector en el que trabaja y que cargo ocupa. Se define si es un usuario profesional orientado a negocios B2B (*business-to-business*) o un usuario común orientado a negocios B2C (*business-to-costumer*)
- **Objetivos y retos:** Se identifican en este apartado cuáles son sus metas, por ejemplo, comprar un producto online, o los retos que ha de superar, como rellenar un formulario.
- **Valores y miedos:** Están relacionados directamente con las motivaciones y objeciones en la consecución de tareas en el sitio web que pueden estar relacionados con el diseño del sitio o con las características del negocio. Debemos identificar aquellos factores que motivan la consecución de objetivos y detectar aquellos puntos de fricción que puede retrasarlos.
- **Uso de las tecnologías:** Nivel de experiencia.
- **Propuesta de valor:** Resumen gráfico de todos los puntos anteriores.

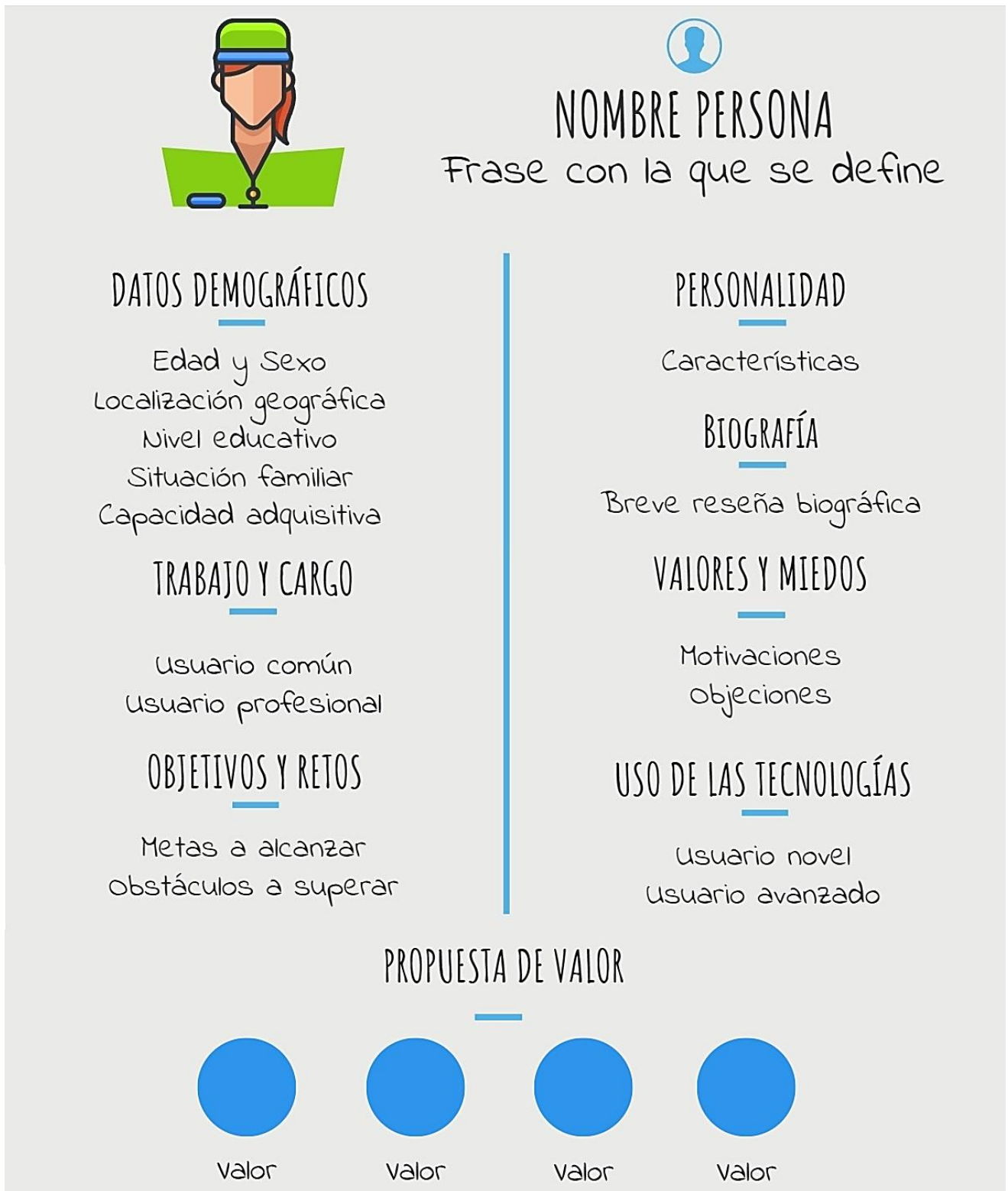


Ilustración 36. Ejemplo de una ficha persona

8.2.2. *User/Customer/Searcher Journey Map*

Entender el comportamiento que el usuario realiza para lograr una meta en el sitio web desde que intuye que tiene una necesidad hasta que logra satisfacerla nos permitirá tener un mayor conocimiento del usuario y contribuirá en la selección de palabras clave, en el diseño de la arquitectura de la información y en el diseño de contenidos. Consta de los siguientes factores clave:

Usuario. Se alinea con la persona y la investigación previa hecha en el análisis y es quien experimenta el viaje.

Scenario y Expectativas. En el Scenario se describe una situación en la que el usuario deberá abordar un objetivo a partir de una necesidad y además se tendrán que definir sus expectativas.

Fases. Son las diferentes etapas que atraviesa un usuario. Dependiendo de la necesidad se generarán distintas etapas, aunque de forma general pueden ser, intención, búsqueda, consideración, acción y recomendación.

Acciones. Se describirán los comportamientos que realiza el usuario.

Motivaciones. Será todo lo relacionado con lo que ilusiona al usuario a la hora de emplear el sitio web y que lo impulsan a avanzar por las siguientes fases del viaje.

Objeciones. Los aspectos desconocidos y miedos frenan el viaje, detectarlas permitirá realizar mejoras en el sitio web o generar contenido que disipe sus dudas.

Oportunidades. Se reflejan los conocimientos adquiridos en la construcción del Journey Map desde el punto de vista de la experiencia de usuario, tanto en cuestiones de diseño y usabilidad como de contenido.



NOMBRE DEL USUARIO

SCENARIO

Organizar un viaje para el próximo verano con mis hijos

EXPECTATIVAS

- Obtener información clara
- Facilidad de pago
- ...

	INTENCIÓN	BÚSQUEDA	CONSIDERACIÓN	DECISIÓN	RECOMENDACIÓN
MOTIVACIONES	Ideas para viajar con niños Visitar Disneyland Paris	Mejores atracciones Disneyland Paris	Ofertas Port Aventura niños gratis	Ofertas hotel Disneyland Paris Ofertas hotel Paris Vuelos baratos Madrid Paris	Opiniones Disneyland Paris
OBJECIONES	Precios hoteles en Disneyland Paris	Disneyland pago a plazos Seguro cancelación de viajes	Qué hacer en Disneyland con niños pequeños		
ACCIONES	Realiza una búsqueda general	Realiza búsquedas más concretas	Pregunta en un foro	Completa una página de pago	Escribir una reseña
OPORTUNIDADES		<ul style="list-style-type: none"> • Contenido • Estrategia • Diseño 	MÉTRICAS	<ul style="list-style-type: none"> • • • 	

Ilustración 37. Ejemplo de *journey map*.

8.2.3. Keyword Research

La selección de palabras clave es un punto clave en cuanto la estrategia de visibilidad del sitio web, además nos aportará un mayor conocimiento del usuario, una mejor arquitectura de la información y un diseño de contenidos enfocado en satisfacer las necesidades del usuario.

Partiendo de la definición de *persona* y del *journey map* deberemos identificar si el sitio web estará orientado a una visibilidad profunda o amplia. Después utilizaremos distintas fuentes para hacer una selección de keywords.

De *persona* y *journey map* se extraen muchas keywords dependiendo de si es una búsqueda informativa o transaccional, si conoce el sitio web y lo que se ofrece en él, si son decisiones impulsivas o reflexivas.

También podemos obtener ideas de palabras clave a partir de las motivaciones y objeciones del usuario, así como de su intención de búsqueda dependiendo de si son del tipo *know*, *do*, *visit web site* o *visit in person*.

Por otro lado, si el sitio no es de nueva creación o ya se disponen de algunos contenidos, también podemos obtener ideas de palabras clave a partir de ellos.

Finalmente, obtendremos *keywords* a partir del estudio de la competencia teniendo en cuenta las *keywords* para las que se posiciona cada uno de los sitios identificados como competidores, las posibles *keywords* que tengamos en común y aquellas que nos diferencien.

8.3. Diseño

8.3.1. Arquitectura de la información

La arquitectura de la información es “el arte y la ciencia de estructurar y clasificar sitios web con el fin de ayudar a los usuarios a encontrar y manejar la información”. (Rosenfeld, 2015)

Sistema de clasificación de la información. Además de tener en cuenta los tipos de clasificación utilizado en usabilidad como son los sistemas exactos y ambiguos, se introduce en este punto los criterios SEO en los que se valora si el contenido es indexable, es decir, que tiene potencial de búsqueda y que tenemos suficiente portfolio, productos, contenidos, etc. que satisfaga al usuario. Si no fuera así, no sólo supone un contenido pobre para el usuario, sino que tampoco será indexado por el buscador.

Estructura de navegación. Los tres tipos básicos de estructuras de navegación pueden ser jerárquico, lineal y en red.

Para su construcción se utilizan métodos como el *Card Sorting* aunque ahora contamos con una gran cantidad de información acerca de cómo busca el usuario y que puede obtenerse a partir del análisis de palabras clave.

Se definirán cuáles serán las categorías principales de navegación, subcategorías y filtros lo que permitirá establecer los niveles jerárquicos y realizar un correcto reparto de popularidad.

Se tendrán en cuenta aquellas búsquedas con mayor potencial y que formarán las categorías principales (por ejemplo, la búsqueda pienso) y las que tengan menos volumen de búsqueda formarán las subcategorías (por ejemplo, pienso para perros). Cuando el usuario utilice en una búsqueda categorías y subcategorías (por ejemplo, pienso para perros esterilizados), está realizando una búsqueda más profunda que puede asignarse a un filtro.

Además, se tendrá en cuenta el **nivel de profundidad** de las páginas del sitio con contenido relevante, es decir a cuantos clics se encuentran de la página principal.

Enlazado interno. Se construirá un buen enlazado interno con url amigables, entendibles para el usuario y para el buscador, y cuyos contenidos sean enlazados desde el menú principal para traspasar autoridad de la portada al resto de páginas que queramos reforzar, así como facilitar el rastreo del robot del buscador.

Errores 404. Aunque no constituye una penalización en cuanto al posicionamiento si lo es para el usuario ya que puede frustrarle si no tiene acceso a contenidos relevantes para él.

Enlaces rotos. En este caso nos fijamos en aquellos enlaces que dirigen al usuario desde nuestro sitio web a otro externo y que por diversas razones no funcionan. Su existencia transmite la idea al buscador de que el sitio no está optimizado y además tiene una imagen negativa para el usuario.

Etiquetado. Las etiquetas no sólo indican al usuario la acción que va a llevar a cabo cuando hace clic sobre ellas, también suponen una fuente de información para el buscador. Se tendrá en cuenta el correcto etiquetado de las siguientes etiquetas y atributos: *title*, *meta description*, encabezados, *alt* y *title* de imágenes, *open graphics*, *lang*, *hreflang*, *meta robots*, *canonical*, *title* de enlace, microdatos estructurados y atributos *rel dofollow*, *nofollow*, *ugc*, *sponsored*.

Sistema de búsqueda del sitio web. La existencia de un sistema de búsquedas en el sitio web permiten al usuario buscar independientemente del menú de navegación y retomar la búsqueda en el sitio si se ha desorientado. Además, se hace necesario incluir un sistema de búsqueda avanzado con una buena gestión de filtros utilizando la información obtenida de las intenciones de búsqueda del usuario. Cuando el usuario utilice en una búsqueda categorías y subcategorías (por ejemplo, pienso para perros esterilizados), está realizando una búsqueda más profunda que puede asignarse a un filtro.

Una vez construida la arquitectura de la información óptima para el usuario se ha de facilitar su navegación al propio robot del buscador. Debemos crear el archivo *robots.txt*, que informará al motor de búsqueda de qué URL debe rastrear o no y el archivo *sitemap* para facilitar la indexación y ahorrar cuota de rastreo.

En definitiva, para una buena arquitectura web será necesario cubrir tres planos, la facilidad de navegación para el usuario, el rastreo por parte de los buscadores y la relevancia interna entre las páginas del sitio web.

8.3.2. Diseño de contenidos

En esta parte se debe tener en cuenta tanto la presentación del contenido debido a las limitaciones de lectura que tiene el uso de pantallas como el contenido en si puesto que es un factor clave en cuanto al posicionamiento.

Brevidad. El contenido debe ser concreto, ha de estar alineado con la intención de búsqueda expresada por el usuario y ha de aportar valor.

Retención. Se busca que el usuario permanezca el mayor tiempo posible en la página o navegando en el sitio web desde el punto de vista de que estamos resolviendo sus necesidades, alimentando sus motivaciones o despejando sus objeciones.

Escaneabilidad. Se trata de facilitar la lectura al usuario que normalmente realiza una lectura rápida para realizar posteriormente una lectura más profunda. Se utilizan las etiquetas de encabezado h para jerarquizar el contenido, el uso de negritas, listas, etc

Lenguaje estructurado. La idea principal ha de situarse en el primer párrafo, no sólo es útil para el usuario, sino que el buscador que se dirige cada vez más hacia una web semántica también lo tiene en cuenta.

Fragmentación. Se trata de ofrecer al usuario la posibilidad de profundizar más en el contenido utilizando textos ancla en los que es recomendable que el *anchor text* coincida con el contenido que estamos enlazando pues se trata de un enlazado interno.

Densidad de palabras clave. Tradicionalmente para posicionar un sitio se hacía un uso excesivo de las palabras clave en el contenido, algo que no aporta nada al usuario y que con el tiempo ha sido penalizado por el buscador considerándolo *Keyword Stuffing*. La densidad de palabra clave no debe ser alta, se han de utilizar sinónimos y ha de buscarse la naturalidad en los contenidos.

Títulos de página. Constituyen la referencia principal del contenido que se presenta en la página y no sólo tienen una gran utilidad para el usuario, sino que tiene una gran relevancia para el buscador. Se definen en la etiqueta *title* y aparecen en las SERPS y es recomendable que contenga la intención de búsqueda.

Snippet en SERPS. Es un resumen del contenido de la página que aparece en los resultados de búsqueda y ha de ser redactada cuidadosamente con un texto atractivo para el usuario. Es recomendable utilizar sinónimos con respecto a la intención de búsqueda de la página.

Legibilidad. El diseñador debe cuidar la legibilidad de los textos, usando tipografías claras, un tamaño de letra adecuado y un buen contraste entre texto y fondo.

Contenido duplicado. Se debe evitar el contenido duplicado tanto en el propio sitio web pues no aportará valor al usuario y el buscador tampoco lo indexará, como fuera del sitio web, ya que el buscador penaliza estas prácticas.

Canibalización. La regla básica es: una intención de búsqueda por página del sitio, mucho más útil para el usuario y concentramos toda la autoridad en una página en lugar de dispersarla.

8.3.3. Diseño visual

En cuanto al diseño visual se realizará teniendo en cuenta su adaptación a los distintos tipos de pantalla prestando especial atención a la versión móvil ya que el buscador construye su índice tomándola como referencia y analiza si su diseño es *responsive* debido a que los usuarios acceden a la web mayoritariamente desde dispositivos móviles lo que también tiene implicaciones con respecto a la velocidad y ancho de banda.

Se debe tener en cuenta que los usuarios no disponen de teclado y ratón lo que hace que debamos prestar más atención a las exigencias de usabilidad además de la incursión de las interfaces de voz que si bien se encuentran todavía en una fase temprana de uso entre los usuarios su tendencia de uso va en aumento.

A la hora de diseñar un sitio web *mobile first* se pueden valorar las siguientes opciones:

- **Diseño responsivo**, donde tendremos un único código fuente que se adapta a los distintos dispositivos.
- **HTML dinámico**, donde tendremos una única URL, pero el código fuente cambia en función del dispositivo.
- **Versión móvil específica.**
- **AMP o Accelerated Mobile Pages**, es un proyecto impulsado por *Google* en el que las páginas utilizan un formato centrado en la velocidad usando un HTML más sencillo de interpretar por los navegadores.

8.4. Prototipo

En la fase de prototipado, cuando tengamos un prototipo de alta fidelidad, se analizará el rendimiento del sitio web ya que constituye, por un lado, un factor clave en el posicionamiento y por otro lado, los usuarios quieren que las webs sean rápidas. Se tendrán en cuenta los siguientes aspectos:

- Elección de un buen hosting
- Uso del protocolo https
- Minimizar el código utilizando JavaScript y CSS en archivos externos
- Utilizar la carga asíncrona retrasando la carga de aquellos elementos que no sean importantes para el usuario en primera instancia.
- Reducir el tamaño de archivos de página, lo que mejorará la velocidad de carga para el usuario y ahorraremos cuota de rastreo para el buscador incrementando la posibilidad de que recorra más contenidos del sitio y los indexe.
- Optimizar el tamaño y peso de las imágenes.
- Implementar un sistema de caché que resuelva los elementos estáticos de manera más rápida.
- Optimizar la base de datos.

8.6. Evaluación

Se deben realiza una evaluación de la usabilidad bien durante las fases de desarrollo del sitio, bien antes de lanzar el producto para detectar errores que sean más fácilmente subsanables antes de llegar a la fase de producción.

Los métodos de evaluación pueden ser realizados con usuarios reales o utilizando métodos heurísticos en los que los usuarios no intervienen.

8.7. Seguimiento

En esta parte se implementan todas las técnicas relacionadas con la popularidad del sitio web, la generación de **backlinks** y estrategias de **Linkbuilding** que están más relacionadas con el SEO y distan completamente de la usabilidad, pues el principal objetivo es ganar posicionamiento en los buscadores al margen de los usuarios.

En esta fase, además, se recomienda implementar una revisión periódica de los principales puntos de la metodología que puede aportarnos información renovada acerca del usuario, ya que no son estáticos y evolucionan a lo largo del tiempo, cambian sus gustos, sus formas de interactuar y se generan nuevas tendencias. Todo ello hace que surjan nuevas oportunidades a la hora de generar contenido o introducir nuevos productos o servicios que puedan llevarnos a modificar partes del modelado del usuario o del diseño o de la arquitectura de la información.

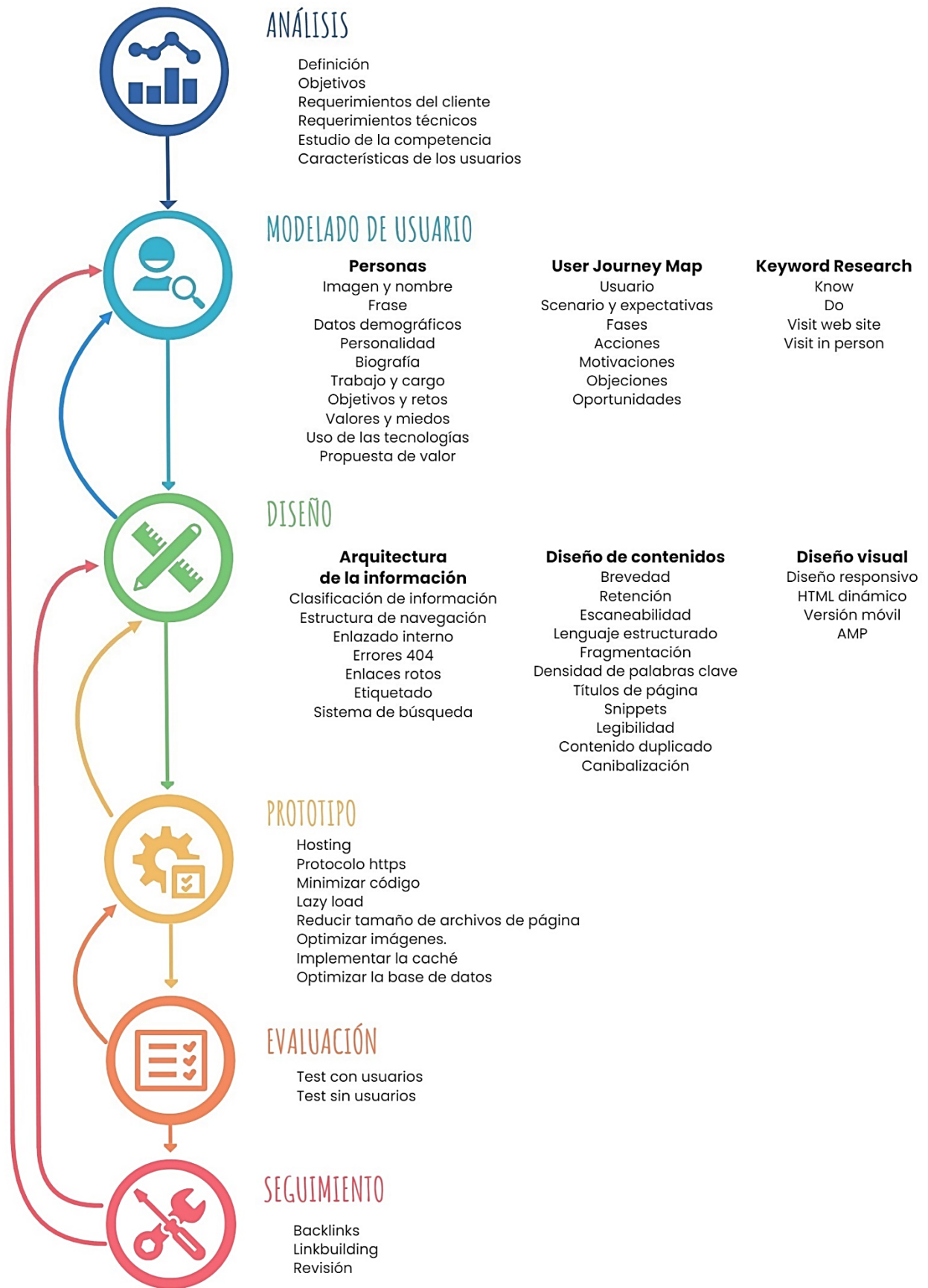


Ilustración 38. Seoabilidad.

8.8. Métricas

Se enumeran en este apartado las métricas que se han de tener en cuenta en el análisis de seobilidad y cuyo cálculo y explicación ha sido ampliamente abordado en el marco teórico:

- Fiabilidad del sitio web
- Número de errores por parte del usuario sean o no involuntarios
- Número de tareas/clics
- Porcentaje de enlaces rotos
- Porcentaje de presencia de propiedades
- Porcentaje de tareas finalizadas con éxito
- Porcentaje de tareas finalizadas con éxito en el primer intento
- Porcentaje de usuarios capaces de finalizar tareas clave sin ayuda
- Tiempo de permanencia
- Tasa de uso
- Tiempo de latencia
- Número de páginas vistas
- Tasa de rebote
- Ticket medio
- Enlaces generados por el usuario de forma natural
- Compartidos en redes sociales
- Comentarios positivos
- Visitas recurrentes
- CTR
- Número de envío de formularios
- Número de conversiones
- Posiciones para las palabras clave del sitio web
- Posiciones por página
- Saturación
- Páginas indexadas por sección
- Número de URL únicas posicionadas
- Número de búsquedas posicionadas

- Número de URL únicas que reciben tráfico orgánico
- URL posicionadas en segunda página
- Core Web Vitals
- URL amigables
- Thin content
- Enlaces internos
- Enlaces totales
- Enlaces follow
- Enlaces no follow
- Dominios únicos de referencia

9. Recomendaciones

En este apartado enumeramos una serie de recomendaciones a tener en cuenta a la hora de aplicar la metodología Seobilidad:

En el análisis

Implicación del cliente. Es básico que el cliente participe tanto en esta fase inicial como en el resto del proceso de diseño, ya que, por un lado, es él quien proporciona las ideas básicas del sitio web y por otro lado, al hacerlo partícipe durante todo el proceso evitamos posibles objeciones.

En el modelado de usuario: Personas

Utilizar datos fehacientes. Debemos alejarnos de nuestras propias percepciones como diseñadores y utilizar investigaciones de mercado o entrevistas con usuarios que sustenten el modelado de usuario.

En el modelado de usuario: User Journey Map

Se debe definir un objetivo concreto. De nuevo debemos dejar los prejuicios a un lado y centrarnos en el usuario. Se partirá de un modelo de persona bien definido y se creará un viaje sencillo y fácil de entender.

En el modelado de usuario: Keyword Research

Si sólo prestamos atención a la **intención de búsqueda** que realiza el usuario en su decisión final y no en todas las pequeñas búsquedas que realiza hasta llegar a la última fase podríamos estar perdiendo muchas palabras clave.

En el diseño: Arquitectura de la información

En algunas ocasiones nos enfrentamos a **cambios en la arquitectura de la información** debido a la falta de stock en productos, productos o contenidos descatalogados y productos o contenidos estacionales.

En el caso de la **falta de stock**, si es una página relevante se debe mantener igual, pero se modificará su contenido incluyendo otros productos que puedan ser de interés.

La gestión de **contenidos descatalogados** se gestiona de distinta forma si tiene potencial de búsqueda o no. En el primer caso, se relega su orden de aparición a final de página dando prioridad a los contenidos más nuevos, y en el segundo caso, se realizaría una redirección 301.

Para los **contenidos estacionales** se dará prioridad de aparición en la página en función de la demanda estacional.

Gestión de filtros. Normalmente no se trabaja su indexación, pero lo cierto es que aquellos que se correspondan con un alto volumen de búsquedas han de ser indexados.

Se debe evitar el uso de páginas **pre-home** a la hora de diseñar el sitio web, son páginas que a priori parece que aportan mucho al usuario pero que desperdician la relevancia del dominio. Siempre que sea posible se incorporará la información de ese tipo de páginas a la página principal.

La **gestión de contenidos paginados** mediante las etiquetas *prev* y *next* ha quedado obsoleta, así que se debe intentar que las categorías y subcategorías muestren sus contenidos en una única página.

En el diseño: Diseño Visual

Constantemente surgen nuevas tendencias en el diseño visual de los sitios web es por esta razón que se deben incorporar con sumo cuidado ya que a veces las tendencias **estéticas** no hacen que haya una mejor usabilidad ni aportan nada al usuario.

En el diseño: Diseño de Contenidos

Se debe prestar atención a la **densidad de palabras clave** y evitar a toda costa el *Keyword Stuffing*.

Se ha de ser cuidadoso con la **canibalización** de contenidos. Es preferible orientar una intención de búsqueda como objetivo de una única página y en el caso de que tengamos varias páginas orientadas a una misma búsqueda debemos usar la etiqueta *rel="canonical"* que nos ayudará a evitar el contenido duplicado. Hay que prestar atención a su uso ya que si no se implementa correctamente podemos obtener efectos contrarios como llegar a desindexar contenido relevante.

En el prototipo

Uso de AJAX. El uso de esta tecnología mejora en gran medida la usabilidad, pero por el contrario dificulta la Indexabilidad así que se ha de buscar el equilibrio.

De entre todas las opciones que hay para implementar el *mobile first* se recomienda el uso de diseños responsivos.

10. Conclusiones

Como conclusión a este TFG de investigación, me gustaría destacar, por un lado, la importancia que tiene trabajar la usabilidad y el SEO de forma conjunta cuando nos enfrentamos al diseño de un sitio web. Distintas fases del DCU como el análisis, el modelado del usuario, la arquitectura de la información y el diseño de contenidos se ven enriquecidas con la incorporación de las estrategias utilizadas en el SEO.

Por otro lado, quería dilucidar si las métricas relacionadas con la usabilidad tenían un peso mayor en el posicionamiento frente a otras métricas que nada tienen que ver con la satisfacción del usuario, es decir, los *backlinks* y la conclusión ha sido que aún hoy por hoy son éstos últimos los que más relevancia tienen para el buscador. Aun así, la matriz de correlaciones mostraba la importancia de las métricas de usabilidad para lograr un buen posicionamiento.

Para llegar a todas estas afirmaciones he querido implementar un modelo matemático utilizado en las disciplinas de *machine learning* de inteligencia artificial, el análisis de regresión múltiple. La razón principal para utilizarlo ha sido la de introducir un análisis científico, más manejable por la cantidad de datos a utilizar y más objetivo que el que en mi opinión se usa tradicionalmente en SEO. En este punto, quiero hacer hincapié que para un correcto uso de estos algoritmos son necesarias muchas más observaciones a las utilizadas y que puede haber otros modelos matemáticos más adecuados para realizar este análisis, aun así la idea es empezar a utilizar este tipo de modelos para validar hipótesis.

Finalmente, y con todas las bases del marco teórico y práctico se ha creado una metodología a aplicar llamada Seobilidad que facilite la aplicación del SEO y la usabilidad en conjunto a la hora de abordar la construcción de un nuevo sitio web.

En definitiva, la intención con este TFG ha sido la de sentar unas bases para crear una metodología conjunta y que con total seguridad podrá ser mejorada y detallada por profesionales de ambas disciplinas. Si además a todo esto le unimos el potencial de los

modelos matemáticos usados en inteligencia artificial para dar veracidad a los análisis, entonces nos encontramos con una idea que en mi opinión merece la pena explorar en profundidad.

Anexo 1. Entregables del proyecto

- **Competidores.** Carpeta que contiene los archivos csv con los datos de dominios competidores, palabras clave en común, palabras clave diferentes, tráfico estimado y *Backlinks* por cada uno de los dominios analizados.
- **Keywords.** Carpeta que contiene los archivos csv extraídos de *Ubersuggest* que contiene datos de palabras clave, volumen, posición, estimación de visitas y dificultad seo para todos los dominios seleccionados del sector mascotas.
- **Serps.** Carpeta que contiene un archivo csv extraído de *Ubersuggest* y que contiene datos de título, url, dominio, visitas, enlaces, compartidas y autoridad de dominio.
- **Análisis Competidores Keywords.** Archivo jupyter con el que se ha realizado el análisis de palabras clave en *Python*.
- **Análisis de regresión lineal múltiple.** Archivo *jupyter* con el que se ha implementado el modelo de regresión lineal múltiple en *Python*.
- **Datos.** Excel con las métricas recabadas para cada uno de los dominios elegidos para el análisis.
- **Matriz de correlación.** Imagen en JPEG de la matriz de correlación de las métricas utilizadas para el análisis.

Anexo 2. Librerías

SciPy es una librería de código abierto para Python que contiene diversas herramientas y algoritmos matemáticos.

Scikit-learn es una librería de software libre que se utiliza con Python que incluye diversos algoritmos de aprendizaje automático.

NumPy es una librería para Python que da soporte al cálculo con vectores y matrices de grandes dimensiones.

Matplotlib es una librería para Python con la que se pueden generar gráficos a partir de datos manejados con *NumPy*.

Pandas es una librería usada en Python que permite la manipulación de tablas numéricas.

Anexo 4. Bibliografía

- Al-Ananbeh, A. &.K. (2012). *Website Usability Evaluation and Search Engine Optimization for Eighty Arab University Websites*. Recuperado el Diciembre de 2021, de https://www.researchgate.net/publication/244990013_Website_Usability_Evaluation_and_Search_Engine_Optimization_for_Eighty_Arab_University_Websites
- Bustamante, E. G. (2021). Mega Guía 2021 para hacer una Auditoría SEO Completa YA. *Aula CM*. Recuperado el Diciembre de 2021, de <https://aulacm.com/guia-hacer-auditoria-seo/>
- Cachón, M. (2021). *Datos de Visibilidad Sectorial - Google Core Update Junio-Julio 2021*. Madrid. Recuperado el Septiembre de 2021, de <http://mjthis.me/core-update-junio-2021>
- Fernández Casado, P. (2021). *UX DESIGN. Hazlo fácil pensando en el usuario*. Madrid: RA-MA Editorial.
- Fontela, A. (2021). Core Web Vitals / Google PageSpeed. *Raiolanetworks*. Recuperado el Diciembre de 2021, de https://raiolanetworks.es/blog/core-web-vitals-google-pagespeed/#conclusiones_sobre_core_web_vitals_y_google_pagespeed
- Gibbons, S. (2018). Journey Mapping. *Grupo Nielsen Norman*. Recuperado el Diciembre de 2021, de <https://www.nngroup.com/articles/journey-mapping-101/?Im=scenario-mapping&pt=youtubevideo>
- Guerra, H. H. (2020). *Aplicación de un modelo de regresión lineal múltiple (MRLM)*. Chile. Recuperado el Noviembre de 2021, de <https://nubeminera.cl/modelo-de-regresion-lineal-multiple-con-python/>
- Harley, A. (2015). Personas Make Users Memorable for Product Team Members. *Nielsen Norman Group*. Recuperado el Diciembre de 2021, de <https://www.nngroup.com/articles/persona/>
- Maciá Domene, F. (2019). *SEO Avanzado, casi todo lo que sé sobre posicionamiento web*. Madrid: Ediciones Anaya Multimedia.
- Maciá, F. (2016). Análisis del viaje del cliente y escenarios de búsqueda. *Human Level*. Recuperado el Diciembre de 2021, de <https://www.humanlevel.com/articulos/posicionamiento-natural-buscadores/analisis-del-viaje-del-cliente-y-escenarios-de-busqueda-1.html>
- McQuade, B. (2020). Definición de los umbrales de las métricas de Core Web Vitals. Recuperado el Diciembre de 2021, de <https://web.dev/i18n/es/defining-core-web-vitals-thresholds/>
- McQuade, B. (2020). Definición de los umbrales de las métricas de Core Web Vitals. *Web.dev*. Recuperado el Diciembre de 2021, de <https://web.dev/i18n/es/defining-core-web-vitals-thresholds/>
- Monjo Palau, T. (2011). *Diseño Centrado en el Usuario*. Barcelona: Eureka Media S.L. Recuperado el Octubre de 2021, de http://cv.uoc.edu/annotation/77847c78a26395a6bb77f8e08b504b8a/485065/PID_00159828/index.html
- Montero, Y. H. (2015). Experiencia de Usuario: Principios y Métodos. En Y. Hassan Montero, *Experiencia de Usuario: Principios y Métodos* (pág. 9). Recuperado el Octubre de 2021, de https://yusef.es/Experiencia_de_Usuario.pdf
- Nielsen, J. (1995). 10 Usability Heuristics for User Interface Design. *nngroup*. Recuperado el Octubre de 2021, de <https://www.nngroup.com/articles/ten-usability-heuristics/>

- Nielsen, J. (2012). Seo and usability. *Nielsen Norman Group*. Recuperado el Diciembre de 2021, de <https://www.nngroup.com/articles/seo-and-usability/>
- Ortega Santamaría, S. (2011). *Introducción a la usabilidad y su evaluación*. Barcelona: Eureka Media S.L. Recuperado el Octubre de 2021, de dfdfdf
- Rodrigo, J. A. (2020). *Regresión lineal con Python*. Recuperado el Noviembre de 2021, de <https://www.cienciadedatos.net/documentos/py10-regresion-lineal-python.html>
- Rosenfeld, L. (2015). *Information Architecture: For the Web and Beyond* (4ª ed.). O'Reilly.
- Tognazzini, B. (2014). *First Principles of Interaction Design* (Revised & Expanded). *AsktoG*, *Interaction designs dolutions for the real world*. Recuperado el Octubre de 2021, de <https://asktog.com/atc/principles-of-interaction-design/>