



SEO semántico : Framework ISS para la optimización de sitios intensivos en contenidos

Carlos Lopezosa¹; Lluís Codina²; Jorge Caldera-Serrano³

Recibido: 8 de junio de 2018 / Aceptado: x de julio de 2018

Resumen: El presente artículo lleva a cabo una revisión sistematizada sobre el concepto SEO semántico y búsquedas semánticas, para configurar un framework de trabajo aplicado a sitios intensivos en contenido. Se pretende crear un procedimiento de trabajo capaz de ayudar a posicionar contenido web en los buscadores de manera enriquecida (semántica). Se revisa de forma teórica los principales elementos que caracterizan el SEO semántico: el gráfico de conocimiento, el algoritmo colibrí, el *RankBrain*, el LSI (Indexación Semántica latente), y los datos estructurados. Seguidamente se analizan las diferentes búsquedas semánticas que se reproducen en índice de Google, y se estudia la iniciativa Schema.org cómo el marcador semántico estandarizado por los principales motores de búsqueda. Finalmente estos elementos se configuran como componentes propios del Framework Integral de SEO Semántico (por sus siglas ISS).

Palabras clave: SEO semántico, entidades, búsquedas semánticas, Google, Optimización en motores de búsqueda, Marcado semántico, sitios intensivos en contenido.

[en] Semantic SEO: ISS Framework for content-intensive web sites optimization

Abstract: This article carries out a systematic review of the concept of semantic SEO and semantic searches, to configure a framework of work applied to content-intensive sites. It is intended to create a working procedure capable of helping to position web content in search engines in an enriched way (semantics). The main elements that characterize semantic SEO are theoretically reviewed: the knowledge graph, the hummingbird algorithm, the RankBrain, the LSI (Latent Semantic Indexation), and the structured data. Next, the different semantic searches that are reproduced in Google's index are analyzed, and the Schema.org initiative is studied as the semantic marker standardized by the main search engines. Finally, these elements are configured as components of the Integral Framework of Semantic SEO (ISS).

Keywords: Semantic SEO, entities, semantic searches, Google, search engine optimization, semantic markup, content intensive sites.

¹ Departamento de Comunicación. Universidad Pompeu Fabra (España)

E-mail: carlos.lopezosa@upf.edu

² Departamento de Comunicación. Universidad Pompeu Fabra (España)

E-mail: lluis.codina@upf.edu

³ Departamento de Información y Comunicación. Universidad de Extremadura (España)

E-mail: jcalsar@unex.es

Sumario: 1. Introducción 2. Objeto, objetivo y metodología 3. Resultados 4. Conclusiones 5. Reconocimientos 6. Bibliografía

Cómo citar: Lopezosa, Carlos; Codina, Lluís; Caldera-Serrano, Jorge (2018). SEO semántico: Framework ISS para la optimización de sitios intensivos en contenidos. *Cuadernos de Documentación Multimedia*, 29, 97-123.

1. Introducción

El objeto de estudio de este trabajo es el concepto de SEO semántico y cómo afecta a los resultados de búsqueda. Se analizan los principales elementos que intervienen en el SEO semántico: Gráfico de conocimiento, algoritmo colibrí, *RankBrain*, LSI, Datos estructurados y Schema.org.

La motivación principal para este trabajo es explorar las posibilidades de desarrollar una propuesta de trabajo, o Framework, de optimización de SEO semántico para sitios intensivos en contenido, para lo cual hemos adoptado ejemplos de sitios vinculados con el sector educativo.

Los objetivos de este informe son:

- Caracterizar el SEO semántico e identificar sus principales componentes.
- Determinar cómo afectan los componentes del SEO semántico a los resultados de búsqueda.
- Proponer un Framework de optimización de SEO semántico que pueda ser aplicable a sitios intensivos en contenidos.

De acuerdo con estos objetivos principales, las preguntas de investigación son las siguientes:

- ¿Es posible identificar el SEO semántico como una nueva tendencia en el posicionamiento web?
- En caso de respuesta afirmativa del punto anterior, ¿es factible caracterizar esta tendencia e identificar sus componentes principales de un modo operativo?
- En caso de respuesta afirmativa del punto anterior, ¿es factible proponer un framework aplicable a sitios intensivos en contenido?

Para poder presentar una respuesta a las tres preguntas anteriores, la metodología utilizada ha consistido en aplicar una revisión sistematizada (Hart, 2008; Booth et al., 2012), de acuerdo con la cual se ha procedido a identificar primero el grupo de fuentes de información cualificadas, después se ha procedido a crear un banco de documentos sobre el tema, como tercer paso se ha llevado cabo una operación de análisis de las referencias seleccionadas y finalmente, hemos procedido a presentar los resultados en forma de la síntesis narrativa crítica que seguirá a continuación y que constituye el grueso de este trabajo.

Para la selección de los documentos hemos utilizado las bases de datos Scopus y Web of Science, cuyos resultados se han completado con el uso de Google Scholar, además hemos utilizado informes y trabajos producidos por organismos y actores de la máxima importancia en este ámbito, por ejemplo, hemos utilizado documentación oficial de buscadores como Google. En el anexo se indican las

ecuaciones utilizadas para las búsquedas en las bases de datos y en Google Scholar y las fuentes adicionales consultadas.

Para el análisis de los ítems que han formado el banco de documentos (ver bibliografía) de esta revisión hemos utilizado este esquema de análisis (Hart, 2008):

- ¿Cuáles son las teorías y conceptos clave en el ámbito considerado?
- ¿Cuáles son las principales bases ontológicas y epistemológicas de la disciplina?
- ¿Cuáles son los temas y debates principales sobre el ámbito de estudio?
- ¿Cuáles son los principales problemas abordados por la disciplina?
- ¿Cómo se estructuran y organizan los conocimientos sobre la disciplina?
- ¿Qué huecos u oportunidades de investigación ofrece el ámbito analizadas

Con este motivo, se realiza un breve recorrido por la historia de este concepto, prestando especial atención a los elementos que lo conforman: gráfico de conocimiento (*knowledge graph*), algoritmo Colibrí, *Rankbrain*, *Latent Semantic Indexing* (LSI), y datos estructurados (Chong, 2017; Google, 2012, 2012a, 2012b, 2018, 2018a, 2018b).

Adicionalmente, se han realizado repetidas pruebas de búsqueda con diferentes palabras clave con el fin de analizar las páginas de resultados, y se ha llevado a cabo un registro sistemático de los diferentes tipos de resultados afectados por el SEO semántico.

Con todo ello pretendemos aportar indicaciones sobre el rol del SEO semántico y cómo afecta a los llamados *rich snippets* o resultados enriquecidos. Recordemos que el *snippet* es como se denomina a los fragmentos de información sobre los sitios web que forman parte de las páginas de resultados.

Por tanto, se denominan *rich snippets* o resultados enriquecidos a los resultados que incorporan elementos adicionales a los resultados estándar. Un resultado enriquecido, por ejemplo, añade iconos o enlaces de navegación al *snippet* convencional.

Como resultado de esta nueva clase de *snippets*, las SERPs (*Search Engine Result Page*) o Páginas de Resultados de los Buscadores están cambiando. Cada vez es más frecuente encontrar resultados de búsqueda más refinados que responden mejor a la intención de búsqueda del usuarios (Codina et al. 2017; Crestodina. 2015).

Estos resultados semánticos están colonizando el así llamado *top 10* (los resultados de la primera página), relegando a un segundo plano los resultados clásicos, es decir, aquellos que no se presentan con componentes enriquecidos.

The image shows a Google search for "blade runner 2049". The search results are organized into several sections, with three specific areas highlighted by red boxes to illustrate semantic search results:

- Top Left Box:** "Cartelera de Blade Runner 2049". It lists showtimes for the movie at various theaters like "Arenas de Barcelona Multiores" and "Palaú Balaña".
- Middle Box:** "Blade Runner 2049 (2017) - Película eCartelera". It provides a link to the eCartelera website and mentions the director Denis Villeneuve and the cast.
- Bottom Left Box:** "Noticias destacadas". It features three news snippets: "El montaje original de Blade Runner 2049 duraba cuatro horas", "China no salta a Blade Runner 2049 en laquilla", and "Blade Runner 2049 iba a durar 4 horas y se dividía en dos partes".

Other visible search results include a "Blade Runner 2049 (2017) - FilmAffinity" entry, a "Blade Runner 2049 (2017) - Película eCartelera" entry, and a "Noticias destacadas" section with three items. The right side of the image shows a detailed view of the movie "Blade Runner 2049" with a poster, a "Reproducir tráiler en YouTube" button, and statistics for FilmAffinity (7,310), eCartelera (7,310), and SenasCine.com (3,95).

Figura 1: Resultados de búsqueda semántica (recuadros rojos) en base a la palabra clave “película blade runner 2049”.

Como resultado las estrategias tradicionales de SEO en sitios intensivos en contenidos ya no son del todo válidas si se desea que los motores de búsqueda muestren sus resultados de manera enriquecida. Como cualquier otra clase de sitio web, necesitan marcar los contenidos de manera semántica para que éstas se posicionen de manera más eficiente en las SERPs.

Ante esto, surgen una serie de retos a los que los sitios intensivos en contenidos deberán enfrentarse en los próximos años para tener mayor visibilidad:

- Por un lado, la web semántica afecta al trabajo del SEO en general. Por tanto, los sitios intensivos en contenido, como una clase más de sitios web, también tendrán que adaptarse también a ellas.
- Por otro lado, la web semántica afecta a las representaciones de los resultados de búsqueda, ofreciendo mayor inmediatez y mayor competencia en el escalado de posiciones en los buscadores.

2. Antecedentes: La web semántica

Tim Berners-Lee, creador de la web y promotor del W3 Consortium (W3C), escribió lo que podría considerarse el ensayo fundacional de la web semántica en mayo de 2001 junto con James Hendler y Ora Lassila (2001).

Por aquel entonces estábamos ante el auge de la web 2.0, una web multidireccional (emisor y receptor podían comunicarse bidireccionalmente) pensada para humanos, pero no para que los ordenadores fueran capaces de entendernos.

Bajo esta premisa, Berners-Lee y el resto de investigadores que publicaron el mencionado trabajo plantearon una web que sería una extensión de la web 2.0 capaz de hacer que un ordenador fuera capaz de entender los contenidos de una forma que imitar el comportamiento inteligente (Codina et al., 2009; Lacy, 2005).

Para lograrlo, teorizaron sobre la posibilidad de usar un conjunto de lenguajes y protocolos estándar que fueran capaces de añadir metadatos que, a su vez, hicieran explícitos los significados de las páginas web y, al mismo tiempo, pudieran ser interpretados por programas de ordenadores.

La web 2.0 que se caracterizaba por ser un medio de interacción entre documentos y personas, y entre personas y personas, gracias a la web social y la computación en la nube, a la web semántica, un medio en el que los ordenadores podrían entenderse entre ellos utilizando el significado marcado y especificado en las páginas web.

De las ideas originales de Berners-Lee y del proyecto de la web semántica del W3C surgen tres elementos fundamentales que caracterizan a esta nueva web: los lenguajes para la descripción de metadatos, las ontologías, y los agentes de usuario.

Lenguajes para la descripción de metadatos

Los metadatos son un conjunto de lenguajes capaces de estructurar y de describir otros datos (p.e, metadatos para describir una página web) o cualquier elemento que tenga característica de entidad, siendo estas últimas cualquier cosa, literalmente, capaz de ser identificada mediante sus propiedades (Codina et al., 2009). De este modo, gracias a los metadatos y al uso de programas capaces de procesarlos, las máquinas pueden usar, gestionar, recuperar y buscar recursos online (Berners-Lee et al., 2001)

De este modo esta clase de lenguajes son uno de los elementos fundamentales de la web semántica. Como hemos señalado, se caracterizan por ser información y datos al mismo tiempo, y se encargan de representar y describir de forma explícita los contenidos de los documentos asociados.

Los lenguajes para la descripción de metadatos esenciales de la web semántica son: XML *-extensible Markup Language-*, RDF *-resources Description Framework-* y OWL *-Lenguaje de Ontologías Web-* (Lacy, 2005).

Probablemente, el modelo de representación de metadatos más utilizado en la web semántica es RDF, sin embargo como este lenguaje tiene ciertas limitaciones suele utilizarse de forma combinada con OWL, con RDF Schema, Skos, y RSS y Atom.

Ontologías, entidades y esquemas

Las ontologías, en el contexto de la web semántica, son esquemas conceptuales para representar entidades del mundo real o simbólico. Estas pueden incluir conceptos y reglas para que los ordenadores sean capaces de realizar razonamientos automáticos a partir de reglas de las reglas de inferencia y de las propiedades y relaciones representada en las ontologías (Codina et al., 2009)

Adicionalmente, las ontologías favorecen la interoperabilidad entre sistemas informáticos con el fin de alcanzar un razonamiento automático, y por tanto conseguir que las máquinas puedan representar el contenido web (Codina et al., 2009).

En el ámbito de la web, el desarrollo principal de las ontologías y otras estructuras relacionadas (como las taxonomías) ha girado en torno al concepto de entidad y, más concretamente, el de tipo de entidad.

Como hemos señalado, las entidades son objetos concretos o abstractos que pueden identificarse y que tienen propiedades. Los tipos de entidad son las clases abstractas que reúnen a entidades que comparten alguna propiedad.

En el caso de los protocolos de trabajo de la web adoptados por los buscadores, el concepto de tipo de entidad se identifica con el concepto de esquema, lo que ha dado lugar a Schema.org que es la estructura del conjunto de esquemas (tipos de entidades) que pueden reconocer los buscadores.

Agentes

Entre el concepto de agente, destaca el de agentes de usuario que son programas informáticos capaz de interpretar la semántica de la web con el fin de poner su contenidos al servicio de las necesidades de información de los usuarios (Codina et al., 2009). Para ello, la web semántica trata de aportar una estructura al contenido significativo. Para ello crean un entorno en el que los agentes puedan realizar tareas sofisticadas para los usuarios.

En el diseño original de la web semántica del W3C, los navegadores deberían ser uno de los agentes de usuario capaces de interpretar la web y llevar a cabo tareas delegadas por los usuarios, por ejemplo, establecer la mejor ruta para un viaje y adquirir los billetes correspondientes según uno de los ejemplos de Berners-Lee, y mucho más no sólo para la revolución digital, sino también para nuestra continua prosperidad, e incluso para nuestra libertad (2010).

3. Qué es el SEO semántico

El término SEO semántico no es nuevo. La primera vez que se utiliza es en 2004. Sin embargo, parece que no es hasta mayo de 2008 que vuelve a utilizarse, al menos de una forma notable (Massimo Paolini, 2017). A partir de 2012, los grandes expertos en SEO comienzan a definir el concepto de SEO semántico sin llegar a un consenso sobre qué es y en qué consiste. Aún a día de hoy, el término no está totalmente consensuado.

Durante 2010 a 2013 se realizan algunas aproximaciones conceptuales muy interesantes, destacando Doszkocs (2010) y Briggs (2012) así como, dado el carácter de este tema, los contenidos de algunos blogs, entre ellos el de un medio de comunicación que se constituyó en auténtica autoridad, denominado Mashable (2012). Esta clase de aportaciones fueron las grandes precursoras en la

conceptualización de las búsquedas semánticas que, años después, derivaría en unas primeras definiciones de SEO semántico como las de Brown (2013) y DiSilvestro (2013) y Paolini (2017).

La idea de SEO semántico comenzó por el concepto de búsqueda semántica (DiSilvestro, 2013).

Tamas Doszkocs (2010) definió las búsquedas semánticas como aquellos resultados producidos incluso cuando los elementos recuperados no contenían ninguno de los términos consultados por el usuario.

Briggs (2012) confirmó que los resultados de búsqueda se estaban moviendo más hacia la recuperación de información basada en entidades en vez de en palabras clave.

Mashable (2012), como ya hemos dicho, uno de los blog de referencia sobre noticias de Internet, también daría su propia definición sobre las búsquedas semánticas, reafirmando que Google utilizaba la inteligencia artificial para entender la intención de las búsquedas y el significado de la consultas.

Brown (2013), confirmó que el SEO del futuro debía de ajustarse al paisaje de la búsqueda semántica y aplicarlo en su trabajo diario.

En la actualidad, el SEO semántico se puede definir, entre otras cosas, como una estrategia para optimizar la verdadera intención de búsqueda de los usuarios. Se trata de un trabajo en donde el SEO debe contestar a una pregunta y luego a otra, y así sucesivamente para conseguir un contenido de mayor valor, y de mayor profundidad. Con ello se consiguen más oportunidades para obtener una variedad de Rankings de palabras clave (Wells Yu, 2017).

A nivel del buscador, Paolini (2017) se basa en la definición de Doszkocs para afirmar que el SEO semántico se refiere a una pregunta o frase de búsqueda que produce resultados significativos incluso cuando los resultados no contienen una coincidencia exacta de las palabras o frases utilizadas. Esto se debe a que el algoritmo de búsqueda es lo suficientemente inteligente como para interpretar la intención del usuario y el contexto potencial de su búsqueda.

Desde el punto de vista de los motores de búsqueda, el SEO semántico consiste en buena parte en adivinar la intención del usuario y entregar las cosas (información) que se considera que está buscando incluso en situaciones de ambigüedad o cuando el usuario no expresa bien su búsqueda.

Desde el punto de vista de productores de contenidos, el SEO semántico se considera como “el arte y la ciencia de adaptar nuestro contenido a los buscadores semánticos” (Kosaka, 2016). Para ello, por ejemplo, el SEO debe enfocarse en temas, y no palabras clave, incluyendo frases semánticamente relacionadas dentro de su contenido.

Por tanto, el SEO semántico tiene como objetivo ayudar a los motores de búsqueda a devolver las consultas más relevantes y refinadas, independientemente de si los términos de búsqueda llevados a cabo por el usuario son obvios o no.

Por su parte, la labor del experto en SEO semántico es la de etiquetar contenido de manera semántica, en base a una serie de códigos (Schubert, 2017). Es decir, debe explicar a los motores de búsqueda el significado de cada uno de los rincones

de la web. Para ello hay que posicionar utilizando marcado semántico y enriqueciendo el contenido del sitio web.

Llevando acciones de SEO semántico no solo se ayuda a los motores de búsqueda a entender mejor los contenidos de una web, sino que también, ayuda a la web a aparecer en los índices de búsqueda de manera enriquecida.

A la vista de lo anterior y considerando que el punto común del proyecto de la web semántica (W3C) y de las propuestas de los buscadores para mejorar los resultados con marcado semántico (Schema.org) es el concepto de entidad, nuestra propuesta de **definición** y caracterización sintética de SEO semántico es la siguiente:

El SEO semántico consiste en el posicionamiento basado en el concepto de entidad, su marcado efectivo por parte de los creadores de contenidos y su identificación eficiente por parte de los sistemas de información, lo que incluye en su capacidad para establecer relaciones entre entidades y sus propiedades.

A partir de la definición anterior, podemos decir que el SEO semántico, por tanto, presenta dimensiones que afectan (1) a los buscadores y (2) a los creadores y gestores de contenidos.

Por parte de los buscadores intenta aportar páginas de resultados con capacidad de satisfacer total o parcialmente necesidades de información a los usuarios sin necesidad de que éstos deban acudir a las fuentes originales.

Para ello, los sistema de información deben identificar tipos y ocurrencias de entidades y sus propiedades y relacionarlas entre ellas imitando una actividad que, en caso de llevarla a cabo un ser humano, sería considerada inteligente.

Por parte de los creadores y gestores de contenidos requiere el uso de marcado semántico para que los agentes de usuario (navegadores, p.e.) y los sistemas de información (buscadores, p.e.) pueden procesar la información en términos semánticos, esto es concretamente, en términos de entidades y sus propiedades.

4. Framework Integral de SEO Semántico

Como sabemos por el marco teórico, el SEO semántico afecta tanto a los creadores de contenidos como a los agentes de usuario y a los sistemas de información, siendo el punto de articulación, los conceptos de entidad y de propiedades de entidad.

El Framework Integral de SEO Semántico (ISS) es un marco de trabajo que se sitúa en el plano de los creadores de contenidos, y que está orientado para optimizar webs intensivas en contenidos de forma eficaz a nivel semántico, con el objetivo de que los contenidos aparezcan de forma enriquecida en las páginas de los resultados de búsqueda.

A continuación, mostramos el Framework ISS en forma de diagrama:

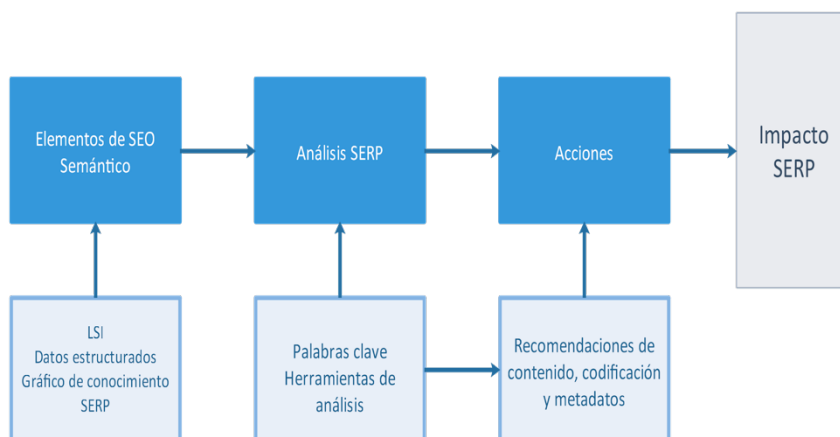


Figura 2: Diagrama del Framework ISS.

El Framework ISS propone enfocar el SEO de forma holística, desde el punto de vista de las múltiples interacciones, de ahí que se justifiquen sus seis componentes que lo integran:

- Gráfico de conocimiento
- Algoritmo Colibrí
- *RankBrain*
- LSI
- Datos estructurados
- Schema.org

El procedimiento general del Framework cuenta además con las siguientes fases:

1. Identificar una o más palabras clave que serán objeto de optimización
2. Llevar a cabo una búsqueda con la palabra clave
3. Analizar los resultados para cada uno de los componentes del Framework, eventualmente, con el uso de herramientas de análisis online.
4. Considerar las medidas de optimización
5. Aplicar las medidas, en su caso
6. Monitorización y nuevas acciones de optimización

A continuación analizamos los elementos que componen el SEO semántico y su aplicación en el Framework.

4.1 Gráfico de conocimiento

El gráfico de conocimiento es una iniciativa de Google que fue lanzada el 16 de mayo de 2012 con la intención de ofrecer resultados de búsqueda interconectados más precisos y relevantes (Google 2012, 2012a).

Se trata de un almacén de conocimiento que recopila datos de personas, empresas, eventos, animales, momentos históricos etc. y nos los devuelve en forma de resultado de búsqueda semántica.

Este banco inmenso de información recoge millones de datos de palabras clave buscadas con frecuencia por los usuarios e interpreta la intención detrás de esas palabras clave, en base al contenido ya existente.

En la actualidad, el Gráfico de Conocimiento muestra información que define a ciertas entidades a partir de datos tomados de multitud de fuentes distintas como la propia Wikipedia, IMDb, Facebook, etc (Hernández, 2014).

El gráfico de conocimiento evalúa estos datos que han sido recogido de miles de millones de búsquedas y determina su relevancia.

De este modo, el Gráfico de conocimiento consigue sus dos principales propósitos:

- Ayudar a Google a refinar sus resultados búsquedas.
- Presentar cuadros de resultados de búsqueda específicos para proporcionar respuestas directas a consultas o temas estrechamente relacionados.

Algunos de los principales responsables del proyecto del Gráfico de conocimiento Jack Menzel, Sashi Thakur, Emily Wright y Johanna Wright daban algunas claves en el vídeo de presentación del mismo:

- A medida que se desarrolle el Gráfico de Conocimiento, Google será capaz de entender mejor las consultas de los usuarios y la información disponible en la web, así como la forma de conectar ambas de una manera inteligente.
- El gráfico de conocimiento ayudará a Google a pasar de motor de información a motor de conocimiento.

El primer experto en dar una respuesta de cómo optimizar el gráfico de conocimiento fue Tony Edward (2015) y su publicación en Search Engine Land. En este artículo ofrece tres pasos para que una web pueda aparecer como gráfico de conocimiento en los resultados de búsqueda.

Estos pasos se centran en:

- Crear un perfil en Wikipedia de la web/empresa
- Crear un perfil en Wikidata de la web/empresa
- Implementar la web con Schema.org en base o bien al esquema *organización* o al esquema *persona*, es decir, a la entidad organización, si es una empresa, o a la entidad *persona*, si es una persona física.
- Configurar adecuadamente el perfil de Google + de la web/empresa

El framework ISS plantea la optimización del gráfico del conocimiento siguiendo las tareas planteadas por Edwards e incluye otras como:

- Incluir marcado localizado y optimizado en Google My business y Google Plus: Google My Business es la plataforma de Google destinada a las empresas. Google Plus es la red social de Google.
- Verificar la web como empresa o como web personal con algún *Plugins* de SEO.
- Verificar las redes sociales de una web con algún *Plugin* de SEO.

4.2 Algoritmo Colibrí y algoritmo Rankbrain

El Algoritmo Colibrí es un algoritmo creado por Google que se lanzó de manera oficial el 20 de agosto de 2013. Colibrí se centra principalmente en las búsquedas semánticas y el gráfico de conocimiento.

No es una actualización, sino que fue un algoritmo nuevo. Su función es la de entender mejor la intención de búsqueda de los usuarios, ofreciéndoles los resultados más relevantes (Pedraza, 2017).

Colibrí cuenta con tres componentes clave:

- Buscar conversaciones: el algoritmo colibrí permite obtener resultados de búsqueda muy refinados que mejoran la conversión entre la búsqueda realizada por el usuario y los resultados óptimos de las SERPs. Colibrí proporciona a los usuarios finales una página de resultados más íntima y personal.
- Búsqueda humana: Cuando realizamos una búsqueda, Colibrí se enfoca en los sinónimos y en los temas relacionados con la temática consultada, lo que ayuda a encontrar mejores resultados de búsqueda.
- Fundamentos de la búsqueda por voz: Gracias a Colibrí los resultados locales se volvieron más precisos; ya que permite combinar el procesamiento conversacional del lenguaje con la comprensión de la intención humana en base a sus datos de localización.

La misión de Colibrí es la de comprender el significado de las búsquedas, para ello el buscador tuvo que pasar de identificar las cadenas de búsqueda a identificar las entidades de búsqueda.

Las entidades de búsqueda hacen referencia a búsquedas personalizadas. Para Google, las palabras no son “cosas” en sí mismas sino la representación verbal de dichas “cosas”. Las entidades de búsqueda son la forma en la que Google objetiva las palabras en conceptos.

Otro de los algoritmos que forman parte del seo semántico es el RankBrain, éste junto con el Gráfico de Conocimiento y el algoritmo colibrí es uno de los grandes elementos de las búsquedas semánticas. RankBrain es una variable más del algoritmo de Google que se hizo pública en agosto de 2015. Utiliza la inteligencia artificial para ofrecer resultados de búsqueda adecuados en base a búsquedas nuevas o ambiguas (González, 2016).

A nivel práctico, RankBrain recibe una palabra o una frase desconocida, adivina con qué palabras o frases podrían tener un significado más adecuado, y lo filtra para ofrecer los mejores resultados de búsqueda, en base a consultas nunca realizadas con anterioridad.

Rankbrain ha supuesto una gran revolución, y se ha convertido en la Inteligencia Artificial que utiliza Google en todos sus resultados de búsqueda (Google, 2012b). Además, controla el algoritmo de Google por completo, ya que, en la actualidad, éste se ha vuelto tan complejo, que no puede controlarse de forma manual (Schachinger, 2015).

No se sabe mucho sobre RankBrain, debido a las restricciones por parte de Google a sus patentes, sin embargo gracias a la publicación parcial de la patente, y a otras fuentes información como Bloomberg (primer medio en informar sobre

RankBrain), o Search Engine Land y Search Engine Journal nos permiten identificar algunas de sus claves (Stetzer, 2015):

- RankBrain forma parte del algoritmo de búsqueda general de Google (no es un algoritmo nuevo) que ayuda a clasificar páginas y encontrar los resultados de búsqueda más refinados para consultas particulares.
- Rankbrain es uno de los cientos de factores que componen el algoritmo de Google, sin embargo, en 2016, Google reveló que es el tercer factor de posicionamiento más importante.
- RankBrain se utiliza para interpretar las búsquedas que los usuarios envía para encontrar páginas que podrían no tener la palabra clave de búsquedas consultadas pero que sí se ajustan de forma eficaz a lo que busca el usuario.
- RankBrain realiza un autoaprendizaje centrado en traducir una búsqueda ambigua en algo más específico, y atraer así mejores respuestas.

Ante este nuevo cambio del SEO tradicional al SEO semántico, los expertos en el campo, tuvieron que empezar a entender cómo funcionaban las entidades de búsqueda.

Esto provocó que pasaran de escribir para los motores de búsqueda en forma de tesoro, a tratar de optimizar las webs en base a las conexiones entre las cosas.

Para ello, resultaba necesario ofrecer a Google un contexto fácilmente comprensible para los temas en torno a los cuales se crease una página web.

Incluimos en el Framework ISS las recomendaciones de Gianluca Fiorelli (2013), publicó en el blog de MOZ algunas tareas que servirían para llevar a cabo una optimización en base a Colibrí y ayudar al entendimiento del contenido multimedia por parte de Rankbrain:

- Utilizar una buena arquitectura de la información.
- Evitar problemas de canonización.
- Evitar contenido de baja calidad (menos de 300 palabras).
- Creación de contenido semántico.
- Uso de lenguaje natural, semánticamente rico.
- Contenido útil utilizando diferentes formatos, uso de redes sociales, vídeos, etc.
- Implementación del contenido con marcado semántico.

4.3 LSI

LSI son las siglas en inglés de Indexación Semántica latente. Como su nombre sugiere, es el análisis de la semántica oculta en un texto. Es un método matemático utilizado para determinar la relación entre términos y conceptos en el contenido (Hurley, 2017).

Los contenidos de una página web son rastreados por un motor de búsqueda y las palabras y frases más comunes son compiladas e identificadas como las palabras clave de la página.

LSI trata de entender los términos, conceptos y todo lo que rodea a un contenido concreto para tratar de entender su contexto y así poder relacionar de manera más



Figura 4. Ejemplo de autocompletado con la palabra clave “cursos online C”. Se puede observar el número de búsquedas mensuales y su coste por clic. En rojo hemos seleccionado una palabra clave que podemos incorporar en el texto para enriquecerlo semánticamente.

A continuación contamos con la herramienta gratuita **Isigraph.com**. Es un programa que nos devuelve palabras clave que podremos utilizar para mejorar nuestro LSI.

Debemos introducir la palabra clave principal a posicionar, y nos devolverá resultados semánticos que después podremos utilizar en nuestro texto (página/entrada).

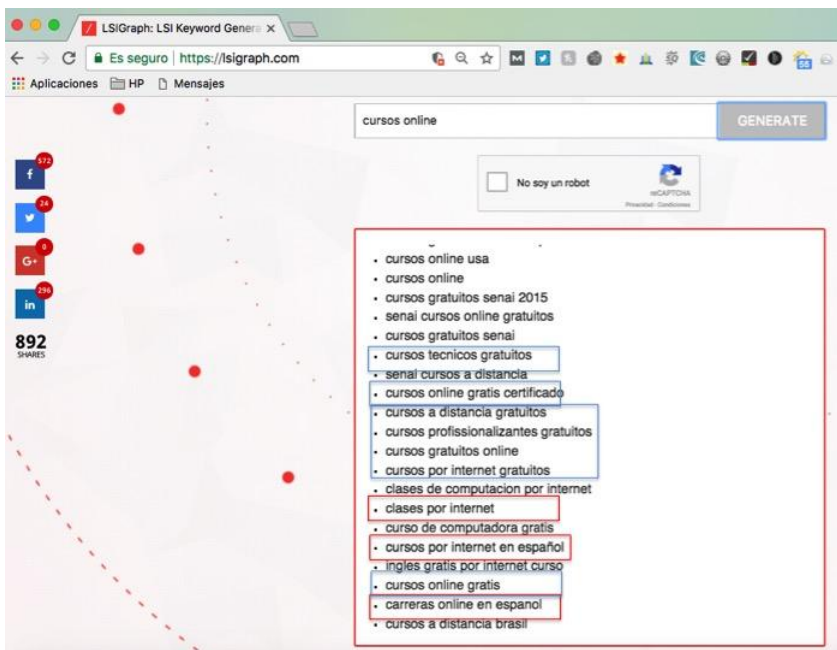


Figura 5. Resultados LSI para la palabra clave “cursos online” utilizando la herramienta <https://isigraph.com>. Podemos incorporar estas palabras en el texto a optimizar para enriquecerlo semánticamente.

Seguidamente, buscaremos resultados enriquecidos en forma de preguntas y respuestas dentro del buscador. Para ello podemos realizar búsquedas de la palabra clave incluyendo “qué”, “Cómo”, “cuándo”, “cuánto”, etc.

Otras preguntas de los usuarios

¿Qué es un curso en línea?

La formación online son **cursos que se realizan de forma no presencial** a través de un dispositivo con conexión a Internet, se utilizan TIC. Los estudiantes pueden trabajar desde su casa o desde cualquier sitio en el que tengan **conexión a Internet** .

Diferencias entre cursos online y cursos a distancia - educaweb.com
www.educaweb.com/...online.../diferencias-formacion-online-formacion-distancia/

Buscar: ¿Qué es un curso en línea?

¿Qué es estudiar online?

Estudiar **online** es una tendencia que ha ido cogiendo fuerza en los últimos años. La oferta formativa de cursos **online** es cada día mayor, tanto en **temática** como en tipo de formación.

¿Por qué estudiar online? - educaweb.com
www.educaweb.com/contenidos/educativos/formacion-online.../razones-estudiar-online/

Buscar: ¿Qué es estudiar online?

¿Cómo son los cursos en línea?

¿Cómo funcionan los cursos online?

1. **Flexibilidad** : es posible acceder a los cursos a cualquier hora y en cualquier lugar.
2. En EXIT Formación te abrimos una puerta a una **nueva experiencia formativa** en modalidad eLearning o Teleformación.
3. Descarga tu solicitud de información de **formación privada** .

¿Cómo funcionan los cursos online? - Exit Formación
exitformacion.com/exit/index.php/77.../87-como-funcionan-los-cursos-online

Buscar: ¿Cómo son los cursos en línea?

¿Qué son los estudios en línea?

¿Cuáles son las ventajas y desventajas de estudiar en línea?

¿Qué es el e learning?

¿Qué es la educación en línea?

¿Qué es estudiar a distancia tradicional?

Figura 6. En rojo las palabras clave que debemos incorporar en cada una de las respuestas que se contestarán en nuestro texto (página/entrada). A medida que desplegamos una pregunta, se despliegan en las SERPs nuevas preguntas.

Una vez veamos las preguntas, debemos desplegar las respuestas, y seleccionar las palabras clave que tendrán que incluirse dentro texto. De este modo tendremos más posibilidades de aparecer en estos resultados enriquecidos.

Una vez tenemos localizadas las palabras clave LSI, y las preguntas y respuestas debemos saber qué otras palabras clave genéricas debemos incorporar y en qué proporción y densidad. Para ello contamos con la herramienta SEOlyze.

SEOlyze es una herramienta de *SEO On Page* que nos ayuda a optimizar nuestros textos en base a una palabra clave a posicionar. Cuanto mejor optimicemos nuestro texto en base a la fórmula de SEOlyze más probabilidades tendremos para gustar a los motores de búsqueda y posicionarnos mejor, más rápido y por más palabras clave relacionadas.

SEOlyze nos ayuda a escribir contenidos optimizados en base al valor WDF - IDF, que es una versión optimizada del valor TF -IDF1 (*Term frequency – Inverse document frequency*).

Es un valor que tiene en cuenta la densidad de la palabra clave principal a posicionar, la densidad del resto de términos relacionados con dicha palabra clave,

y también tiene en cuenta todas las palabras de contexto que se utilizan para posicionar. Para ello SEOlyze tiene en cuenta los 10 primeros resultados de Google. Tenemos que ir al apartado “palabra clave” e incluir el término por el que queremos posicionarnos.

Figura 7. El resultado que se nos da es una media de los diez primeros resultados de Google en base a la palabra clave “cursos online”.

Lo que encontramos en Seolyze en base a los diez primeros resultados:

- **Word count:** La media de palabras utilizadas por el top 10 para posicionarse por “cursos online”.
- **H1-Encabezamientos:** Porcentaje de veces que se usa la palabra clave a posicionar como h1.
- **H2-Encabezamientos:** Porcentaje de veces que se usa la palabra clave a posicionar como h2.
- **H3-Encabezamientos:** Porcentaje de veces que se usa la palabra clave a posicionar como h3.
- **Strong-Tags:** Porcentaje de veces que se usa la palabra clave en negrita, o como etiqueta.
- **Title-Tag:** Se enumera cuántas webs del TOP 10 usan la palabra clave a posicionar en la etiqueta título.
- **Meta-Description:** Se enumera cuántas webs del TOP 10 usan la palabra clave a posicionar en la metadescripción.
- **Keyword in Title:** Se enumera cuántas webs del TOP 10 usan la palabra clave a posicionar en el título.
- **Keyword in Description:** Se enumera cuántas webs del TOP 10 usan la palabra clave posicionar en la descripción.



Figura 8. Gráfica con las palabras que se usan para posicionar “cursos online” en base a una tres palabras.

La gráfica nos muestra varios elementos adicionales:

- **Palabras:** Los términos, que aparecen en la parte inferior de la gráfica, son las palabras que se utilizan en los textos del top 10 para posicionarse por la palabra clave “cursos Online”, por lo que si queremos posicionarnos por esta palabra clave tendremos que incluir todas esos términos en nuestra página/entrada.
- **Gráfico naranja intenso:** esta gráfica muestra las veces que se repite cada término a lo largo del texto, de modo que cuanto más alto es el gráfico más veces se repite un término.
- **Gráfico naranja suave:** este gráfico muestra los picos máximos que no podemos superar de repetición de los términos.

Hay que incluir todos los términos que da SEOlyze y conseguir una gráfica que esté por encima del naranja intenso, pero por debajo del naranja suave. De este modo los textos estarán mejor optimizados y gustarán más a Google ya que contarán con todas las palabras de contexto necesarias para posicionar para una palabra clave y sus derivadas semánticas.

4.4 Datos estructurados y schema.org

Los datos estructurados son informaciones enfocadas a máquinas que consisten en etiquetas que describen el contenido de una página web. Estas etiquetas no pueden ser vistas por los usuarios, sin embargo, los robots sí pueden verlas e interpretarlas.

Cuando implementamos los datos estructurados en una web estamos proporcionando, a los buscadores, información sobre su contenido en forma de conceptos y significados (Enge, 2015). Esto ha provocado que los buscadores hayan evolucionado.

Hemos pasado de obtener resultados de búsqueda en donde los buscadores debían recuperar la información en base a una palabra clave, a contar con elementos semánticos, que permiten una mayor interpretación del contenido y más en concreto a la intención de búsqueda. Esto favorece la obtención de resultados en las SERPs mucho más refinados (Rovira et al., 2013).

Cuando los buscadores son capaces de comprender el contenido de las páginas web, éstos pueden habilitar un conjunto de características para sus páginas de resultados (Meyers, 2016).

Las características de búsqueda se implementan utilizando datos estructurados de dos modos. El primero, agregando información de manera que los motores de búsqueda sean capaces de entender en qué consiste una web; y el segundo agregando mejoras a dicho contenido.

Con ello se obtienen dos tipos de resultados de búsqueda semántica generales (Google, 2018, 2018b):

- **Búsquedas semánticas en base a las características de contenido:** Son resultados de búsqueda semánticos relacionados con el tema de la web. Por ejemplo, si tenemos una web de recetas, los buscadores nos pueden dar un resultado en forma de receta.
- **Mejoras:** Son resultados de búsqueda mejorados que se dan cuando agregamos mejoras o características adicionales a los datos estructurados de características de contenidos. Por ejemplo, cuando incluimos estrellas de valoración en una web, los buscadores devolverán resultados en donde se mostrarán las estrellas valorativas.

Estos son los principales resultados de búsqueda semánticos en base a las mejoras adaptadas a los datos estructurados de los sitios web.

Para que los motores de búsqueda sean capaces de ofrecer este tipo de resultados semánticos, contamos con Schema.org. Se trata de una de las iniciativas más destacadas del movimiento semántico creadas por Google, Yahoo y Bing en 2011 (meses más tarde se uniría Yandex).

Schema.org es una ontología o esquema conceptual muy preciso que estandariza y simplifica el proceso de marcado web.

Esta iniciativa muestra una colección de esquemas de metadatos, o vocabularios (Sule, 2015) que se utilizan para definir, de manera codificada, las propiedades de una web y hacerla más comprensible para los buscadores.

Dicho de otro modo, es una colección de palabras que cualquiera puede usar para describir sus contenidos web. Schema.org no solo favorece a los buscadores sino también a las páginas web marcadas.

Al codificar una web con Schema.org los motores de búsqueda lo reconocen, lo entienden y seleccionan los datos más pertinentes para mostrarlos en los resultados

de búsqueda de forma enriquecida. Con ello se consigue destacar frente a los resultados clásicos (Tort; Olivé, 2015).

Schema.org cuenta con su propia web de referencia y en ella podemos ver cada uno de los esquemas con los que podemos marcar una web. A su vez, cada esquema se divide en entidades y propiedades.

La entidad es el concepto que se quiere definir o representar, y las propiedades las características del concepto. Las entidades están organizadas jerárquicamente de modo que se subcategorizan, además cada entidad, independientemente de su jerarquía, puede heredar propiedades de las entidades jerárquicamente superiores (Nogales et al., 2015)

Cada propiedad cuenta únicamente con el tipo de dato que puede contener y este puede ser: un valor, una enumeración y una entidad.

El éxito de Schema.org radica en que ha conseguido simplificar y estandarizar el proceso de marcar semánticamente el contenido web, al mismo tiempo que proporciona beneficios por realizar este marcado, en forma de fragmentos enriquecidos.

Para implementar los datos estructurados en un sitio web necesitamos llevar a cabo un proceso de etiquetado o marcado semántico consistente en describir el contenido de nuestra web siguiendo como guía las ontologías de Schema.org (Yandex, 2018).

Hay diferentes maneras de optimizar los datos estructurados, sin embargo las tres más conocidas son (Google, 2018a): Microdatos, RDFa y JSON- LD

Dependiendo del tipo de dato que usemos deberemos realizar un tipo de proceso u otro.

Microdatos y RDFa

Los Microdatos y RDFa son un conjunto de etiquetas HTML que tienen como objetivo asociar el contenido de una página con Schema.org (Schema.org, 2018).

Para poder llevar a cabo el marcado semántico con Microdatos y RDFa tendremos que intercalar las descripciones en el contenido (Bing, 2011).

A continuación se muestra un ejemplo de marcado con Microdatos. Se trata de una página que muestra el nombre de una película en este caso **Blade Runner 2049**, información sobre el director, género de la película y trailer:

```
<div itemscope itemtype = "http://schema.org/Movie">
  <h1 itemprop="name">Blade Runner 2049</h1>
  <span>Director: <span itemprop="director"> Denis Villeneuve </span> (born
October 03, 1967 )</span>
  <span itemprop="genre">Science fiction</span>
  <a href=" ../movies/blade-runner2049-trailer.html"
itemprop="trailer">Trailer</a>
</div>
```

Las etiquetas en negrita describen las características conocidas de la película mediante atributos:

- **Itemscope**: Indica que se va a hablar de alguna cosa dentro del bloque <div>.
- **Itemtype** – Indica que ese algo es una película.
- **Itemprop** – Indica las características conocidas de la película (nombre, director, género, URL del trailer).

JSON-LD

JSON-LD es probablemente el marcado semántico que se acabe convirtiendo en el más utilizado ya que es el que Google recomienda. JSON-LD etiqueta semánticamente el contenido incrustando la descripción en fragmentos de código separado (script). De esta manera la implementación es más sencilla, ya que el código se puede situar en bloque sobre cualquier zona de la página a marcar (Schema.org, 2018).

Para poder llevar a cabo el marcado semántico con JSON-LD se le añade un bloque de código con todas las descripciones. A continuación se muestra un ejemplo de marcado con JSON-LD. Se trata de una página que muestra el nombre de un actor, en este caso Harrison Ford:

```
<script type="application/ld+json">
{
"@context": "http://schema.org",
"@type": "Person",
"name": "Harrison ford",
"image": "http://harrison-ford.org/gallery/harrison-ford.jpg",
"url": "http://harrison-ford.org"
}
</script>
```

El método sigue un orden lógico. Cada elemento que describe las etiquetas permite una implementación muy clara. Sin embargo, la mayor pega de utilizar JSON-LD estriba en que se debe repetir el proceso para cada una de las páginas que van a marcarse semánticamente.

Para poder optimizar nuestros datos estructurados, necesitamos marcar nuestro contenido con schema.org, para ello podemos realizar el marcado semántico. Tenemos las siguientes posibilidades:

- A mano: Marcar semánticamente el contenido de una web a mano supone ir a la página de schema.org y coger el esquema que se relacione con una web, rellenar los datos y pegarlos en el código HTML de la web en cuestión.
- Con herramientas: Marcar semánticamente una web con herramientas es un método que consiste en rellenar un formulario con nuestros datos y una herramienta los transformará en microdata, RDFa y JSON-LD según nos interese.

- Con Plugins: Marcar contenido con plugins es la mejor opción si utilizamos un gestor de contenidos web como pudiera ser un Wordpress, Joomla, Drupal, Magento, etc.
- Con marcadores de datos: Esta forma de marcado es la mejor en caso de que no tengamos acceso interno a la web. Es una alternativa virtual al marcado de datos web que permite el etiquetado de los datos de manera externa sobre la página mediante una herramienta que proporcionan los propios buscadores.

4. 5 Optimización de los resultados de búsqueda

El último paso del Framework ISS hace referencia a la mejora de nuestros fragmentos enriquecidos que aparecen en los resultados de búsqueda. Para ello contaremos con la herramienta SISTRIX, y más concretamente con su módulo, SERP-Snippets.

SISTRIX Toolbox es una de las herramientas de análisis SEO más potentes y famosas del mercado. Ha ganado numerosos premios y se utiliza en más de 30 países. Entre algunas de sus funcionalidades destaca su índice de visibilidad, su ranking de palabras clave, su análisis de construcción de enlaces, etcétera.

Para el Framework ISS contamos con el módulo de SERP-Snippets. Entendemos como SERP-Snippets los títulos y descripciones de las páginas que aparecen en los resultados de búsqueda.

SERP-Snippet	Palabra clave	Posición	Características SERP	Tráfico	Competencia
<p>Licenciatura en Publicidad y Relaciones Públicas - Facultad de ... - UPF</p> <p>https://www.upf.edu/facom/es/titulacions/publicitat/p-rppp-ll/</p> <p>Presentación. La licenciatura en Publicidad y Relaciones Públicas de la UPF está orientada a la formación de profesionales con un alto nivel de rigor y ...</p>	licenciado relacione...	1	€ [icon] [icon] [icon] [icon] [icon] [icon] [icon]	[Progress bar]	[Progress bar]
<p>Grado en Biología Humana - Grados (UPF)</p> <p>https://www.upf.edu/es/web/graus/grau-biologia-humana</p> <p>El grado en Biología Humana pretende formar biólogos capacitados para incorporarse en la actividad profesional dentro del campo de la biomedicina y de la ...</p>	biologia humana	1	€ [icon] [icon] [icon] [icon] [icon] [icon] [icon]	[Progress bar]	[Progress bar]
<p>Sonetos, por orden cronológico - Todo Góngora - (UPF)</p> <p>https://www.upf.edu/todogongora/poesia/sonetos/</p> <p>Sonetos, por orden cronológico. Sobre dos urnas de cristal labradas - De pura honestidad templo sagrado -</p>	sonetos de luis de g...	1	€ [icon] [icon] [icon] [icon] [icon] [icon] [icon]	[Progress bar]	[Progress bar]

Figura 9. Interface de SISTRIX SERP-Snippets y resultados en base a la web upf.edu.

Para mejorar los Snippets con el módulo SERP-Snippets de SISTRIX debemos utilizar los filtros que aparecen en la plataforma.

- **Título de Snippet se corta:** Si incluimos este filtro, aparecerán todos los SERP-Snippets que tienen el título cortado por el buscador.

- **Texto de Snippet se corta:** Si incluimos este filtro, veremos todos los SERP-Snippets con una metadescripción demasiado extensa, y por tanto el buscador la recorta.
- **No se muestra el texto:** Si incluimos este filtro, aparecerán los resultados en los que no se ha llevado una labor de inclusión de títulos y metadescripciones.
- **Se muestra tan sólo una línea:** Si incluimos este filtro, aparecerán todos los SERP-Snippets en los que podemos incluir más contenido ya que el título o la metadescripción se han quedado cortas.

Por lo tanto se deben analizar cada una de los resultados de SERP- Snippets, en base al apartado “SERP-Features”. Con este módulo controlaremos e identificaremos cómo aparecen las páginas de una web en los índices de resultados, para mejorar sus fragmentos enriquecidos.

5. Aspectos destacados y conclusiones

El SEO semántico se está imponiendo porque ayuda a las webs a aparecer en las páginas de los resultados de búsqueda (SERP) de una forma mucho más atractiva. Por su parte, RankBrain e ILS son actores esenciales para obtener resultados en la búsqueda semántica. Así mismo, Schema.org es el paradigma del SEO semántico. Es un lenguaje estandarizado y elegido por los principales motores de búsqueda para favorecer los resultados enriquecidos.

En conjunto, la estructura y composición de las SERP están siendo cada vez más afectadas por el SEO semántico, de modo que parece ser una tendencia muy estable y que irá aumentando cada vez más, puesto que otros factores empujan en la misma dirección como las búsquedas por voz en el móvil y en dispositivos domésticos.

Google y otros buscadores, notablemente, Bing, reconocen cada vez más tipos de entidades, así como están cada vez en mejores condiciones de establecer relaciones semánticas entre tales entidades. Es previsible que, cada vez, las SERP se orienten a ser sistemas de respuestas directas, en lugar de ser sistemas de recuperación de información.

Para presentar las conclusiones tomaremos las preguntas de investigación a fin de considerar su grado de cumplimiento.

Pregunta 1 ¿Es posible identificar el SEO semántico como una nueva tendencia en el posicionamiento web?

La riqueza y amplitud de los resultados obtenidos permiten confirmar de forma positiva esta pregunta. Cabe considerar que el término se conoce desde el 2008 y, aunque se avanzó de una manera mucho más profunda en su conceptualización durante 2012, el SEO semántico comienza a cobrar protagonismo durante 2016.

Cada vez es más habitual encontrar artículos y reportajes sobre SEO semántico realizados por profesionales del sector, empleados de Google y medios digitales centrados en el SEO y el Marketing Online.

Pregunta 2. En caso de respuesta afirmativa del punto anterior, ¿es factible caracterizar esta tendencia e identificar sus componentes principales de un modo operativo?

De acuerdo con nuestros resultados, sí, ya que si identificamos cómo funcionan los elementos esenciales que conforman las búsquedas semánticas, en este caso, el gráfico de conocimiento, Algoritmo colibrí, *Rankbrain*, LSI, datos estructurados y schema.org podremos plantear estrategias SEO alineadas con todos estos elementos.

Para ello se plantea una estrategia que hemos llamado Framework ISS, que optimiza un contenido en base a una palabra clave teniendo en cuenta la optimización LSI, la Optimización de Datos estructurados, la optimización del gráfico de conocimiento, y la optimización de las SERPs.

Pregunta 3. En caso de respuesta afirmativa del punto anterior, ¿es factible proponer un framework aplicable a sitios intensivos en contenido?

Nuestros resultados identifican que sí, ya que cuanto mejor sea nuestro contenido (más palabras clave relacionadas, sinónimas y semánticamente asociadas entre sí), y mejor marcado semántico en base al uso de schema.org, mejor entenderá Google nuestro contenido y por tanto mayores posibilidades de posicionarnos en el índice de resultados.

Tener un proceso de trabajo que tenga en cuenta los elementos que utiliza Google sobre las búsquedas semánticas, nos ayudará a alinearnos de manera más eficaz en el proceso que tiene Google de entender los resultados de búsqueda, por lo que seremos premiados copando más y mejores posiciones en las SERPs.

Aportaciones destacadas este trabajo

Por último, nos gustaría resaltar las que consideramos que son las principales aportaciones de nuestra investigación:

- Hemos contribuido a clarificar el concepto de SEO semántico, señalando conceptos, componentes, y procedimientos.
- Hemos aportado un Framework de optimización del SEO semántico.

7. Referencias

- Alberto Nogales, Miguel-Angel Sicilia, Salvador Sánchez- Alonso, Elena Garcia-Barriocanal (2015). Linking from Schema.org microdata to the Web of Linked Data: An empirical assessment, *Computer Standards & Interfaces// Computer Standards & Interfaces*. 45 (Marzo 2016) 90-99. doi: 10.1016/j.csi.2015.12.003
- Berners-Lee, T.; Hendler, J.; Lassila, O. The semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities// *Scientific American*, (2001). 01-03.
- Berners-Lee. Long live the web//*Scientific American*, (Dic. 2010).80-85.
- Bing. (2011) Marking up your site: Overview. <https://www.bing.com/webmaster/help/marking-up-your-site-with-structured-data-3a93e731> (2018-01-15).

- Booth, Andrew; Papaionnou; Sutton, Anthea (2012). *Systematic Approaches to a Successful Literature Review*. London: Sage, 2012.
- Brown, Matthew. (2013). 5 Questions About Semantic SEO. <https://moz.com/blog/semantic-seo-questions> (2018-01-17).
- Briggs, Justin. (2012) Entity Search Results. <https://www.briggsby.com/entity-search-results-the-on-going-evolution-of-search/> (2018-01-17).
- Chong, Joanne. (2017). The Ultimate Semantic SEO Guide: How To Do SEO Like An Expert. <https://seopressor.com/blog/semantic-seo-guide-how-to-do-seo/> (2018-01-22).
- Codina L, Gonzalo-Penela C, Pedraza-Jiménez R, Rovira C. (2017) Posicionamiento web y medios de comunicación ciclo de vida de una campaña y factores SEO. Barcelona: Serie Digital Digidoc. Proyecto Comunicación Interactiva, 2017.
- Codina L, Iglesias-García M, Pedraza R, García-Carretero L. (2016) Search engine optimization and online journalism: the SEO-WCP framework. Barcelona: Serie Editorial DigiDoc. Active audiences and journalism, 2016.
- Codina, Lluís; Mari-Carmen, Marcos; Pedraza, Rafael (2009). *Web semántica y sistemas de información*. Ediciones TREA, 2009.
- Crestodina, Andy. (2015). How to Future-Proof Your Search Engine Marketing: 5 Tips for Semantic SEO. <https://www.orbitmedia.com/blog/semantic-seo/> (2018-01-22).
- DiSilvestro; Amanda. (2013) Semantic Search Means Semantic CEO: What It Means for you. <https://www.highervisibility.com/blog/semantic-search-means-semantic-seo-what-it-means-for-you/> (2018-01-30).
- Doszkocs, Tamas. (2010) Semantic a search engines mean well // *Online*, 34(2010) 36-42.
- Edward, Tony. (2015). Leveraging Wikidata To Gain A Google Knowledge Graph Result. <https://searchengineland.com/leveraging-wikidata-gain-google-knowledge-graph-result-219706> (2018-02-07).
- Enge, Eric. (2015). The Definitive Guide to Google's Rich Answers. <https://www.stonetemple.com/the-growth-of-rich-answers-in-googles-search-results/> (2018-02-30).
- Fiorelli, Gianluca. (2013). Hummingbird Unleashed. <https://moz.com/blog/hummingbird-unleashed> (2018-02-07).
- Giomelakis, Dimitrios; Veglis, Andreas. (2015). Employing Search Engine Optimization Techniques in Online News Articles// *Studies in Media and Communication*, ISSN 2325-8071. 3:1 (June, 2015) 22–33. doi:10.11114/smc.v3i1.683
- Giomelakis, Dimitrios; Veglis, Andreas. (2016) Investigating Search Engine Optimization Factors in Media Websites// *Digital Journalism*. 4:3 (2016) 379-400. DOI: 10.1080/21670811.2015.1046992
- González, Juan. (2016). El Futuro del SEO: ¿Cómo afectarán RankBrain y las búsquedas por voz? <http://useo.es/búsquedas-voz-rank-brain/> (2018-02-13).

- Google (2018). A reintroduction to Google's featured snippets. <https://www.blog.google/products/search/reintroduction-googles-featured-snippets/> (2018-02-14).
- Google. (2018a) Introduction to Structured Data. <https://developers.google.com/search/docs/guides/intro-structured-data#structured-data-guidelines> (2018-02-14).
- Google. (2018b) Search Gallery. <https://developers.google.com/search/docs/guides/search-gallery> (2018-02-14).
- Google. (2012). Introducing the Knowledge Graph: things, not strings. <https://googleblog.blogspot.com.es/2012/05/introducing-knowledge-graph-things-not.html> (2018-02-15).
- Google. (2012a). Pagina web oficial del Gráfico de Conocimiento de Google. (2018-02-15). <https://www.google.es/intl/es/insidesearch/features/search/knowledge.html>
- Google. (2012b) Patente: Using concepts as contexts for query term substitutions. <https://www.google.com/patents/US9104750> (2018-02-16).
- Hart, Chris (2008). *Doing a Literature Review: Releasing the Social Science Research Imagination*. London: Sage, 2008.
- Hernández, Chris. (2014). El uso de Knowledge Graph en Google. <https://www.adrenalina.es/knowledge-graph-google/> (2018-01-22).
- Hurley, Sam. (2017). Smashing Semantic SEO in 2017 & Beyond: The Ultimate Guide. <https://www.digitalcurrent.com/seo-engine-optimization/semantic-seo-guide/> (2018-01-30).
- Kosaka, Kim. (2016). Semantic SEO: How to Change Your Game to Win in Search. <https://blog.alex.com/semantic-seo-how-to-change-your-game-to-win-in-search/> (2018-02-13).
- Lacy, Lee W. (2005). *OWL: Representing Information Using the web Ontology Language*. Canadá: Ed. Trafford, 2005.
- Mashable. (2012). How Google's Semantic Search Will Change SEO. <https://mashable.com/2012/03/22/google-semantic-search-seo/#whzognAiIGqV> (2018-02-14).
- Meyers, Peter. (2016). Posic Ranking #0: SEO for Answers. <https://moz.com/blog/ranking-zero-seo-for-answers> (2018-01-30).
- Nastic, Roko. (2010). What Is Latent Semantic Indexing. <https://www.searchenginejournal.com/what-is-latent-semantic-indexing-seo-defined/21642/> (2018-01-22).
- Paolini, Massimo. (2017). What is Semantic SEO? <https://online-sales-marketing.com/what-is-semantic-seo/> (2018-01-22).
- Pedraza, Beau. (2017). How the Google Hummingbird Update Changed Search <https://www.searchenginejournal.com/google-algorithm-history/hummingbird-update/> (2018-01-30).
- Rovira, Cristòfol; Codina, Lluís; Monistrol, Ricard (2013). Rich snippets: información semántica para la mejora de la identidad digital y el SEO // El profesional de la información, 22:1 (2013) 554-561. <http://dx.doi.org/10.3145/epi.2013.nov.08>

- Schachinger, Kristine. (2015). How RankBrain Changes Entity Search. <https://searchengineland.com/rankbrain-changes-entity-search-234345> (2018-02-13).
- Schema.org. (2018) Documentation. <http://schema.org/docs/documents.html> (2018-02-27).
- Schema.org. (2018a) Organization of Schemas. <http://schema.org/docs/schemas.html> (2018-02-27).
- Schubert, Kathrin. (2017). Semantic SEO Guide <https://en.ryte.com/magazine/semantic-seo-guide> (2018-01-23).
- Stetzer, Adam. (2015). RankBrain: SEO friend or foe? <https://searchenginewatch.com/sew/opinion/2434672/rankbrain-seo-friend-or-foe> (2018-02-13).
- Sulé Duesa, Andreu (2015). Schema.org, la mejora de la visualización de los resultados en los buscadores y mucho más. // *BiD textos universitaris de biblioteconomia i documentació*, 34 (Junio 2015) 1-7.
- Tort, Albert; Olivé, Antoni Olivé (2015). An approach to website schema.org design. // *Data & Knowledge Engineering table of contents archive*. ISSN: 0169-023X. 99, (2015) 3-16. doi>10.1016/j.datak.2015.06.011
- Yandex. (2018) Schema.org. <https://yandex.com/support/webmaster/schema-org/what-is-schema-org.xml> (2018-02-27).
- Yu, Wells. (2017). How to Use Semantic SEO for Higher Rankings <https://www.searchenginejournal.com/content-semantic-seo/201596/> (2018-01-23).

8. Anexo

A continuación, se muestra una tabla con las bases de datos consultadas, fuentes de autoridad utilizadas y algunos ejemplos de ecuaciones de búsqueda llevadas a cabo durante la investigación. Las informaciones utilizadas en concreto para el trabajo se pueden ver en la Bibliografía.

Bases de datos	Ejemplo de ecuaciones	Fuentes de autoridad
Scopus	Semantic web, Schema.org, “SEO” AND “Semantic web”, Semantic SEO, “SEO” AND “web”, “Search engine” AND “web”, y diversas variaciones de idiomas y sinónimos.	Google
Web of Science		Bing
ACM		Yandex
Google Scholar		Search Engine Journal
		Search Engine Watch
		Search Engine Land
		MOZ
		Alexa