

Aprovecha el SEO Semántico y optimiza realmente tu web

SEO semántico. En algunos círculos ya se habla mucho de esto, pero en general se le tiene muy olvidado en el mundo hispano. Y a veces cuando se habla de el se dicen un montón de cosas que la verdad ni tienen que ver o se confunden conceptos...

Pero bueno, aquí les diré yo **qué es el SEO semántico**, que conceptos hay que aprender a manejar y como aprovecharlo para optimizar en serio su contenido 😊

Todo esto viene de la evolución de la web a una web semántica. No es nada nuevo, se propuso desde 1998! Por el mismísimo **Tim Berners-Lee**. Fue propuesta como una extensión a la web actual, una extensión que fuera más allá de enlazar con hipervínculos un montón de documentos. Con esta web semántica, **se pretende que en lugar de enlazar documentos se enlace información**, apoyándose en el uso de entidades (identificadas por **URIs** globales y únicas) y basándose en relaciones llamadas tripletas. Todos estos conceptos los manejarás al terminar el artículo 😊

Primero: Qué es semántica

Antes de comenzar a hablar que se entiende por SEO Semántico y como aprovecharlo... hay que definir que es semántica.

La semántica tiene que ver con el significado de las palabras y de las relaciones entre estas.

Así, sencillo. Si lo quieren complicar un poco podemos decir que se refiere a

conceptos como el sentido entre frases y palabras, entre signos, su connotación, sinónimos, antónimos y todo lo que en conjunto significan.

Pasamos ahora a la web semántica:

La Web Semántica es una Web extendida, dotada de mayor significado en la que cualquier usuario en Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida. Esta Web extendida y basada en el significado, se apoya en lenguajes universales que resuelven los problemas ocasionados por una Web carente de semántica en la que, en ocasiones, el acceso a la información se convierte en una tarea difícil y frustrante.

— [W3C](#)

Todo esto de la semántica tiene que ver con varios conceptos complicados que probablemente muchos de los que se dedican al SEO jamás entenderán.

Conceptos como:

- **RDF** – Resource Description Framework
- **NLP** – Natural Language Processing
- **NER** – Named Entity Recognition
- **Topic Modeling**

Y varias cosas más. Afortunadamente no hay que entenderlas todas para poder aprovechar esta tecnología a nuestro favor.

Lo que sí hay que entender son un par de conceptos básicos: **Entidades y Tripletas**.

Entidades y tripletas – La base de la

web semántica

Ya no se trata de palabras clave, se trata ahora de entidades, que son cosas básicamente, cualquier cosa, **desde personas y lugares hasta objetos y eventos**. Estas entidades pueden tener múltiples relaciones con otras entidades.

Por ejemplo, en lugar de pensar en Clint Eastwood como keyword, podemos pensar en él como una entidad.

En este caso, **sabemos muchas cosas de esa entidad...** por ejemplo:

- Ha ganado un par de premios oscar, en 1992 y 2004
- Nació en San Francisco, California
- Es fundador de Malpaso Productions
- Es director de películas como Unforgiven y Million Dollar Baby

¿Y esto de qué sirve?

Probemos:



The image shows a Google search interface. The search bar contains the text "quien es el director de unforgiven". Below the search bar, there are navigation tabs for "Web", "Noticias", "Videos", "Imágenes", "Más", and "Herramientas de búsqueda". The search results show "Cerca de 416,000 resultados (0.40 segundos)". The first result is a knowledge panel for "Clint Eastwood", with the subtitle "Unforgiven, Director". To the right of the text is a black and white photograph of Clint Eastwood wearing a dark coat. Below the knowledge panel, there is a short Wikipedia-style description: "Clinton «Clint» Eastwood Jr. es un actor, director, productor, guionista, músico y compositor estadounidense. Su primer papel relevante fue como secundario en el reparto de la serie de televisión Rawhide. [Wikipedia](#)".



¿Cómo lo supo? Esa información aparece en muchos lados: DBpedia, IMDB, Wikipedia... Y de ahí obtuvo esa información. **Sabe que es una entidad y que dirigió esa película.** No buscando una página echa para ganar dinero con AdSense que tuviera un artículo llamado “¿Quién es el director de Unforgiven?”. Ni siquiera Yahoo Answers.

Y aquí viene el otro concepto: Todas esas relaciones (llamados también ‘hechos’, facts) que acabo de enumerar en esa lista **son llamadas tripletas.** Y tienen un formato en común:

Sujeto + Predicado + Objeto

Desglosándolo: Entidad/Nodo + Propiedad/Relación + Valor de la propiedad/Relación

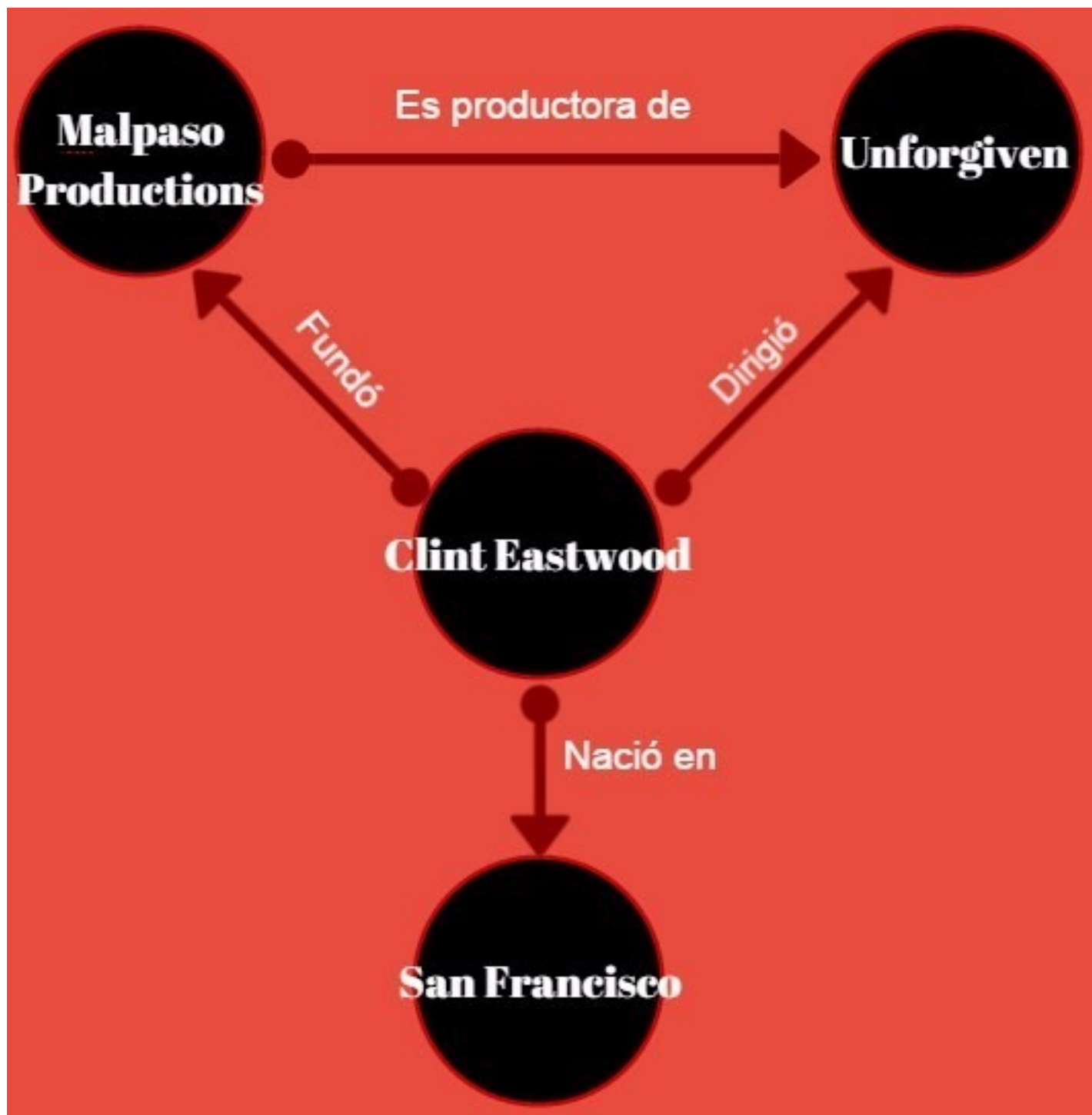
El objetivo de estas tripletas es *representar conocimiento*. Ya no solo datos, sino conocimiento, hechos. Esto implica que cada tripleta tiene un significado, que está dado por el predicado. Entonces, **el sujeto es la entidad que se está describiendo**, **el predicado es qué se está describiendo del sujeto** y **el objeto es el valor del predicado**:

- **Clint Eastwood** **nació** **en San Francisco**
- **Clint Eastwood** **fundó** **Malpaso Productions**

Tanto el sujeto como el predicado y el objeto son entidades, concretas o abstractas, que representan algo en la vida real. Los nombres con que se designa a cada entidad suelen ser globales y son identificados mediante una **URI** (Y digo suelen porque también pueden ser valores ‘literales’ y en blanco,

pero para que nos complicamos).

Estas tripletas se pueden representar perfectamente en grafos. Siguiendo el ejemplo de Clint:



Psst... Hablando de grafos... ¿Conocen el **Knowledge Graph**? De manera general **es una base de datos gigantesca de entidades y relaciones entre ellas**. Y la actualización de Google que usa mucho este Knowledge Graph es el Colibrí (**Hummingbird**), aunque no es tan famoso como **Google Panda** o **Penguin**.

De hecho **esa actualización tiene mucho que ver con la semántica.**

básicamente toma búsquedas largas (generalmente hechas desde móviles, “conversacionales”) y complejas y las deconstruye, buscando conceptos, desambiguando entidades y sus relaciones entre ellas para devolver un resultado adecuado.

Por ejemplo, al buscar “¿Cuál es el mejor restaurante de comida árabe en el DF?” Google recibe esa oración y la mapea con entidades conocidas. Así sabe que debe de devolver (ubicaciones de restaurantes), específicamente de qué tipo (comida árabe) y en qué ubicación (DF).

Esto arroja resultados muy similares si busco “Mejor restaurante de comida libanesa en la capital de México”. **Sabe que el significado es el mismo.**

Aprovechando todo esto: SEO Semántico

El SEO Semántico tiene que ver con ayudarle a los buscadores a entender exactamente de qué tratan tus paginas, mediante diversas técnicas que ahorita vamos a ver. Y no, **no se trata solo de meter sinónimos y ya.**

Para empezar, hay muchos otros tipos de relaciones entre palabras aparte de los sinónimos/antónimos: **Hiperónimos, Hipónimos, Holónimos...**

Y además, no es tan sencillo como meter algunos sinónimos y ya. Si en verdad quieren optimizar lo más posible alguna página, hay una serie de pasos a seguir:

1. Determinar las entidades correspondientes a la página
2. Desambiguarlas directamente
3. Desambiguarlas indirectamente

Vamos a ver una por una:

Determinar las entidades correspondientes a la página

Lo primero que hay que hacer es identificar las entidades de la página.

¿Cómo? Hay varias herramientas que pueden usar:

- [AlchemyApi](#)
- [Semantria](#)
- [TextRazor](#)

Pero aquí hay un detalle... para el Español no funcionan tan bien, funcionan mucho mejor cuando el texto está en Inglés. Para mostrarles un ejemplo, voy a retomar a Clint Eastwood y analizando las entradas en Wikipedia tanto en inglés como en español con AlchemyApi:

En inglés:

Entities	Mount Suribachi	France	Eastwood	American Sniper	Gran Torino	Piedmont High School	Korean War	United States Army	American Film Institute	Time magazine	Madison County	Josely Wales	Richard Schickel	New York Post
Keywords	Best Director	Sergio Leone						lifetime achievement						
Taxonomy														
Concepts	Universal Studios	Academy Awards		Alcatraz	Melrose	CBS	Best Director and Producer	Acher Winsten	Pale Rider	California	Roger Ebert	Eastwood		
Document Sentiment								James Garner	San Francisco					
Targeted Sentiment	Sergio Leone	director		Dirty Harry	New York City	Sondra Locke		Venice Film Festival	The New York Times	Clint Eastwood	Eastwood			
Relations	Harry Callahan	Clinton Eastwood		Don Siegel				Steven Spielberg	Janet Maslin					
Language								United States						
Title	producer			America				Frank Rich	United Artists					
Author														
Text														
Feeds	Entity	Relevance	Sentiment	Type	Subtypes				Linked Data					
Microformats	Clinton Eastwood	0.866307	mixed	Person	Actor Composer MusicalArtist Politician AwardNominee AwardWinner Dedicatee Dedicator FilmActor FilmDirector FilmMusicContributor FilmProducer MilitaryPerson MusicalGroupMember TVActor TVDirector				dbpedia freebase yago					

Entidades

Entities	⊕ Clinton "Clint" Eastwood, jr. born
Keywords	⊕ Clinton "Clint" Eastwood, Jr. (born May 31, 1930) is an American actor, film director, producer, musician, and poli...
Taxonomy	⊕ These roles have made him an enduring cultural icon of masculinity.[1] [2]
Concepts	⊕ These roles, among others, have made him
Document Sentiment	⊕ Eastwood won Academy Awards
Targeted Sentiment	⊕ Eastwood receiving nominations
Relations	⊕ Eastwood directing many of his own star vehicles
Language	⊕ Eastwood has directed films
Title	⊕ Eastwood has also directed films
Author	⊕ he did not appear, such as the mystery drama Mystic River (2003) and the war film Letters from Iwo Jima (2006), ...
Text	⊕ The war drama biopic American Sniper (2014) set box office records
Feeds	⊕ He received considerable critical praise in France
Microformats	⊕ some which were not well received in the United States including several films
	⊕ several films were not well received in the United States
	⊕ several films received in the United States
	⊕ He been awarded two of France's highest honors
	⊕ He has been awarded two of France's highest honors
	⊕ he became a recipient of the Commandeur of the Ordre des Arts et des Lettres and in 2007 he was awarded the...
	⊕ Eastwood has run his own production company

Relaciones

En Español:

Entities	Entity	Relevance	Sentiment	Type	Subtypes	Linked Data
Keywords	Eastwood	0.881008		Person		
Taxonomy	Clint Eastwood	0.624357		Person		
Concepts	Harry Callahan	0.558405		Person		
Document Sentiment	The New York Times	0.49786		PrintMedia		
Targeted Sentiment	Arthur Knight del	0.490606		Person		
Relations	Sergio Leone	0.458547		Person		
Language	Unidos	0.416883		Continent		
Title	San Francisco	0.413464		City		
Author	Los	0.407717		Person		
Text	Francia	0.40714		Country		
Feeds	mejor película	0.401244		EntertainmentAward		
Microformats	Harry	0.401188		Person		
	Don Siegel	0.399416		Person		

Entidades

Entities	This feature is not currently supported for this language
Keywords	
Taxonomy	
Concepts	
Document Sentiment	
Targeted Sentiment	
Relations	

Hmm... ni quería

Así que si quieres nuevas ideas sobre entidades y relaciones para poder incluirlas en tu texto... prueba analizando alguna página autoritativa sobre tu tema en inglés (como Wikipedia).

Ya que tienes una idea de las entidades hay que desambiguarlas, indicando información única a cada una. Esto se puede hacer de forma directa o indirecta, vamos por la primera:

Desambiguar directamente las entidades

La forma más directa de indicarle a los buscadores de que entidades estamos hablando en una página es marcándolas. De preferencia con **Schema**.

Schema es un esfuerzo colaborativo entre **Google, Microsoft, Yahoo y Yandex** para definir un vocabulario estándar con el cual se puedan marcar entidades y sus relaciones. Se puede implementar usando RDFa, Microdata y JSON LD.

Los primeros dos, RDFa y Microdata, consisten en modificar el HTML directamente. Poniendo un ejemplo básico con un **restaurante**:

```
Restaurante PepeYToño
222 Paseo de la Reforma
Cuauhtémoc, Ciudad de México 94086
(408) 714-1489
<a href="http://www.greatfood.com">www.greatfood.com</a>
Horario:
Lunes-Viernes 11:00am - 6:00pm
```

Texto "normal"

```
<div itemscope itemtype="http://schema.org/Restaurant">
  <span itemprop="name">Restaurante PepeYToño</span>
  <div itemprop="address" itemscope itemtype="http://schema.org/PostalAddress">
    <span itemprop="streetAddress">222 Paseo de la Reforma</span>
    <span itemprop="addressLocality">Cuauhtémoc</span>,
    <span itemprop="addressRegion">Ciudad de México</span> <span itemprop="postalCode">12345</span>
  </div>
  <span itemprop="telephone">(123) 456-7890</span>
  <a itemprop="url" href="http://www.greatfood.com">www.greatfood.com</a>
  Horario
  <meta itemprop="openingHours" content="Mo-Fr 11:00-18:00">Lunes-Viernes 11:00am - 6:00pm
</div>
```

Texto con microdata

```
<div vocab="http://schema.org/" typeof="Restaurant">
  <span property="name">Restaurante PepeYToño</span>
  <div property="address" typeof="PostalAddress">
    <span property="streetAddress">222 Paseo de la Reforma</span>
    <span property="addressLocality">Cuauhtémoc</span>,
    <span property="addressRegion">Ciudad de México</span> <span property="postalCode">12345</span>
  </div>
  <span property="telephone">(123) 456-7890</span>
  <a property="url" href="http://www.restaurantepyt.com">www.restaurantepyt.com</a>
  Hours: <meta property="openingHours" content="Mo-Fr 11:00-18:00">Lunes-Viernes 11:00am - 6:00pm
</div>
```

Texto con RDFa

Para esto puedes basarte directamente en la página de Schema para ver que propiedades puedes poner. O puedes apoyarte de algún generador de

Schemas como: [Schema Creator](#) (o su [plugin para WordPress](#)). Incluso puedes usar la herramienta de marcado de datos de Google, que se encuentra [aquí](#) dentro de la Search Console (Google Webmaster Tools).

Y finalmente mi favorita, **JSON LD**. Es una forma alternativa para añadir Schema a una página web... **sin editar directamente el HTML**. Y esto es bastante útil, antes era un poco engorroso estar marcando el HTML con las propiedades correctas. Con JSON-LD es bastante sencillo, ya no hay que editar el contenido, simplemente hay que añadir un `<script>` dentro del HTML y ya!

Para que puedan ver la diferencia, y la potencia de JSON-LD, aquí un ejemplo:

```
<script type="application/ld+json">
{
  "@context": "http://schema.org",
  "@type": "Restaurant",
  "address": {
    "@type": "PostalAddress",
    "addressLocality": "Cuauhtémoc",
    "addressRegion": "Ciudad de México",
    "postalCode": "12345",
    "streetAddress": "222 Paseo de la Reforma"
  },
  "name": "Restaurante PepeYToño",
  "openingHours": [
    "Mo-Fr 11:00-18:00"
  ],
  "telephone": "(123) 456-7890",
  "url": "http://www.restaurantepyt.com"
}
</script>
```

Esto dice lo mismo que los códigos anteriores... pero no hay que manipular el HTML de una página. Solo insertar ese código en la página y listo 😊

Algo muy importante es que **hay que marcar TODO lo que podamos**. Por ejemplo, podríamos agregar también el logotipo oficial del restaurante y sus redes sociales:

```
"telephone": "(123) 456-7890",  
"url": "http://www.restaurantepyt.com",  
"logo": "http://www.restaurantepyt.com/imagenes/logo-pyt.png",  
"sameAs" : [ "http://www.facebook.com/restaurantepyt",  
             "http://www.twitter.com/restaurantepyt",  
             "http://plus.google.com/+restaurantepyt" ]  
}  
</script>
```

Y si ya trabajando en esto te das cuenta que hay una entidad en particular que aún no existe en Schema... siempre puedes **extenderla**.

Antes de pasar a la otra forma de desambiguar entidades, falta algo muy importante: **Comprobar que el código que vamos a incluir para nombrar y describir a estas entidades sea valido**. Para eso hay varias herramientas:

- **La oficial de Google**. O usando **este bookmarklet**.
- **La oficial de Yandex**.
- **JSON-LD Playground**. Específicamente para JSON LD.
- Hay también extensiones para los navegadores. Yo he probado algunas para Chrome y ninguna me ha servido completamente... así que si alguien conoce alguna, que comente! 😊
 - **Semantic Inspector**
 - **Microdata.reveal**

Desambiguar indirectamente las entidades

En última instancia podemos tratar de desambiguarlas indirectamente, mediante nuestro contenido. **Las palabras que usamos dentro de una página pueden enviarle a los buscadores señales acerca del contenido de tu página, lo cual influencia como la interpretan.**

Y como ya dije, no se trata solo de poner algunos sinónimos y ya. Se trata de hacerle ver a los buscadores que tu página trata de cierto tema.

Por ejemplo, si estamos hablando de “ejercicio” no basta con añadir sinónimos como “entrenamiento” y “acondicionamiento físico”. Hay que darle un contexto a esta información. O como dice [David Harry](#) en [este artículo en Search Engine Journal](#), hay que crear “**cubetas semánticas**”.

Por ejemplo, si estamos hablando de como crear una rutina para el gimnasio, podemos tener una cubeta con términos como:

- Ejercicio
- Entrenamiento
- Gimnasio
- Cardio
- Press
- Tonificar
- Pesas
- Ejercicios para abdomen
- Masa muscular
- Proteína
- Fitness
- Rutinas
- Suplementos
- Crossfit
- Squats

- Peso Muerto
- Actividad Física

Y estos términos, hay que incluirlos en nuestro contenido. Hay que hacer que estos términos “co-ocurrán” en el texto, así la daremos señales más fuertes a los buscadores acerca de la temática de nuestro contenido.

Co-ocurrencia no solo se refiere a que aparezcan en el mismo artículo, también importa la frecuencia con la que aparezcan juntos y la proximidad con la que lo hagan.

¿Y de qué sirve esto? En base al análisis de millones de documentos, los buscadores determinan que hay ciertos términos que suelen co-ocurrir en aquellos documentos que hablan de cierto tema. Y si en nuestro texto nosotros también los incluimos, le damos señales al buscador de que nuestro artículo también trata de eso. **Esto sirve para desambiguar entidades.**

Por ejemplo, si ven los buscadores este artículo: **Panda**, ¿De qué podría estar hablando? Podría estar hablando de un animal, de un antivirus y hasta de un grupo de música... Pero en este caso, los buscadores “ven” que dentro del texto co-ocurren palabras como:

- Web
- Google
- SEO
- Contenido
- Autoridad

Y con cierta seguridad pueden decir que en realidad se trata de “Panda”, la actualización de Google.

Si en lugar de eso ven **esta otra página**, encontrarían palabras como:

- Oso
- Animales
- Especies
- Bambú

Y sabrían que se trata de un “Panda” completamente diferente.

Por último, **no hay que confundir co-ocurrencia con co-citación** Para entender co-citación rápidamente:

- Sitio A menciona (con o sin enlace) al sitio B.
- Sitio A menciona al sitio C.
- El sitio B y C son co-citados por el sitio A y se crea una ‘relación’ entre ellos.

Conclusiones

- Aprovechando el **SEO Semántico pueden mejorarse las posiciones y obtener más tráfico**, al lograr que los buscadores reconozcan mejor los temas que tratan las páginas de tu sitio.
- Una de las ventajas más inmediatas de marcar el código con Schema, aparte de desambiguar entidades, son los **Rich Snippets**. Aquí pueden ver una [presentación de Dr. Pete con 85 ejemplos](#)! Ese tipo de resultados mejora el CTR directamente.
- Si tengo un artículo hablando de Hidalgo... ¿Estoy hablando del estado de Hidalgo? O de Miguel Hidalgo, o de la película Hidalgo... **Esto lo puedo definir fácilmente con Schema**. El primer caso correspondería a **State**. El segundo a **Person**. Y el tercero, a **Movie**. Y marcando los datos con sus Schemas correspondientes, **desambiguamos la entidad** Hidalgo y le indicamos a los buscadores de qué estamos hablando.

Pero... hay que tener en cuenta que muchos de los cambios que anuncian

Google y otros buscadores aplican más para el .com, para el mercado anglosajón (A excepción de Yandex, Baidu y compañía). En idiomas mucho más 'complicados' como el Español **tardan más en implementar todas estas mejoras.**

Por más que digan que la web de hoy YA es semántica, todavía falta. Y falta más en el mercado hispano. Puedes ver varios ejemplos de eso en este post de Sergio Redondo: [Web semántica: Google, tú puedes hacerlo mejor](#). Vamos, busquen Sala Bagdad jaja 😊

Eso sí, **esto no es algo que dentro de unos años vaya a desaparecer**. Es el siguiente paso lógico de la web actual, un paso que ya se está dando. Y no solo Google con su Knowledge Graph (y Vault), también Bing con su [Bing Snapshot](#).

Esto va más allá de si crees que eres [Black Hat o White Hat](#).

Ya para cerrar esto, les dejo aquí algunos artículos para aprender más sobre el tema:

EN ESPAÑOL

[Qué son las entidades y su implicación en el SEO](#)

[SEO Semántico: Entiende a los buscadores para que ellos te entiendan a ti](#)

[¿Qué es la búsqueda semántica y por qué me debe importar?](#)

EN INGLÉS

[What is RDF and what is it good for?](#)

[Semantic University](#)

[What is Linked Data? \(Video\)](#)

Implementing JSON-LD in WordPress

Semantic Web

TAGS:

Conceptos Sobre SEO

Tecnicas Recomendadas