

# Buscadores Web y Redes Semánticas

Raquel Rodríguez García  
Universidad Carlos III de Madrid  
Avda. Universidad, 30  
Leganés, Madrid 28911

100063093@alumnos.uc3m.es

María Luisa Zayas de Diego  
Universidad Carlos III de Madrid  
Avda. Universidad, 30  
Leganés, Madrid 28911

100063138@alumnos.uc3m.es

## RESUMEN

En éste trabajo abordaremos el tema de las Redes Semánticas aplicadas a los buscadores que utilizamos tan habitualmente y de cómo se están desarrollando nuevas formas de programación para poder realizar búsquedas más precisas.

## Términos Generales

Web, Semántica, búsqueda, ontología, conocimiento.

## Palabras Clave

Freebase, Hakia, OntoSem, OntoParser, TMR, QDEX y SemanticRank

## 1. INTRODUCCION

Actualmente los buscadores web que nos podemos encontrar son en su gran mayoría buscadores por palabras clave y estadísticas de popularidad. En éste artículo queremos mostrar como poco a poco se van abriendo camino otro tipo de buscadores que trabajan con redes semánticas y ontologías. No es una tarea fácil, ya que requiere mayor almacenamiento de memoria en comparación con otros buscadores, pero para el usuario es mucho más efectivo este tipo de búsquedas, y por tanto la tendencia a desarrollar buscadores semánticos irá creciendo a lo largo del tiempo.

Más adelante nos centraremos en Hakia, para poder explicar con un ejemplo cómo funciona un buscador semántico por dentro.

## 2. BUSCADORES CONVENCIONALES ACTUALES

Hoy en día las páginas web utilizan motores de búsqueda que utilizan palabras clave o árboles jerárquicos por temas; el resultado de la búsqueda es un listado de direcciones Web en los que se mencionan temas relacionados con las palabras clave buscadas. Se pueden clasificar en dos tipos:

### 2.1. Índices temáticos

Son sistemas de búsqueda por temas o categorías jerarquizados (aunque también suelen incluir sistemas de búsqueda por palabras clave). Se trata de bases de datos de direcciones Web elaboradas "manualmente", es decir, hay personas que se encargan de asignar cada página web a una categoría o tema determinado.

## 2.2. Motores de búsqueda

Son sistemas de búsqueda por palabras clave. Son bases de datos que incorporan automáticamente páginas web mediante "robots" de búsqueda en la red.

Dos de las desventajas más importantes de estos buscadores son:

1. La escasa precisión o relevancia en los resultados. Se devuelven muchos documentos poco relevantes para la búsqueda ya que la presencia de una palabra clave en un documento no implica necesariamente que éste sea relevante.
2. La excesiva sensibilidad al vocabulario empleado en las búsquedas y por tanto la imposibilidad de obtener a la primera todos los resultados relevantes disponibles. Muchos documentos de interés pueden no incluir palabras clave pero sí, sinónimos o hiperónimos de ellas.

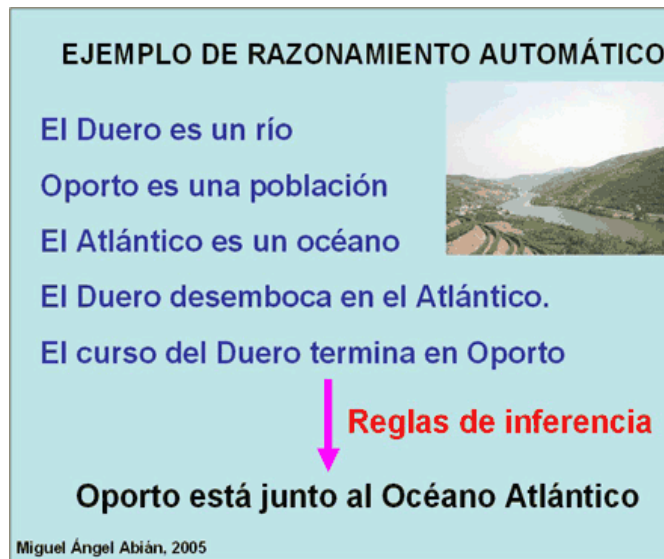
Según un estudio de David Hawking, el porcentaje de páginas relevantes devueltas por un buscador convencional no supera el 0.5%, y esto es lo que intentan solucionar los buscadores semánticos.

## 3. ONTOLOGÍA

Etimológicamente la ontología se puede definir como el logos o conocimiento del ente. En el campo de la informática, utilizamos el término "ontología" para referirnos a la conceptualización formal de un dominio. Normalmente, las ontologías especifican las clases de objetos que existen dentro del dominio en cuestión, las relaciones entre dichas clases, las posibles relaciones entre casos concretos ubicables dentro de una misma clase, y las restricciones aplicables a estos casos. Las ontologías también definen *términos* que denotan estas clases y relaciones, además de objetos individuales.

Las ontologías resultan muy útiles para facilitar el razonamiento automático, es decir, sin intervención humana. Partiendo de unas reglas de inferencia, un motor de razonamiento puede usar los datos de las ontologías para inferir conclusiones de ellos. Por ejemplo, si establecemos estas reglas: "Todos los ríos desembocan

en un mar, en un océano o en un lago" y "Si el curso de un río termina en una población, esa población está junto al mar, océano o lago donde desemboca", las máquinas pueden hacer deducciones como la mostrada en la siguiente figura.



Las ontologías serán imprescindibles en la Web semántica y en los futuros sistemas de gestión empresarial porque permitirán que las aplicaciones estén de acuerdo en los términos que usan cuando se comunican. Mediante ellas, será mucho más fácil recuperar información relacionada temáticamente, aun cuando no existan enlaces directos entre las páginas web. Por ejemplo, una ontología puede usarse para especificar que los buitres son aves carroñeras. De este modo, un buscador que use esa ontología mostrará páginas web sobre buitres cuando un usuario busque información sobre las aves carroñeras.

## 4. REDES SEMÁNTICAS

Una red semántica es una forma de representar el conocimiento lingüístico en la que los conceptos y sus interrelaciones se representan mediante un gráfico. Se utilizan, entre otras cosas, para representar mapas conceptuales y mentales.

En una red semántica se representan esos elementos semánticos mediante nodos. Cuando hay una relación semántica entre dos elementos semánticos se representará en la red semántica como una línea, flecha o enlace entre ambos.

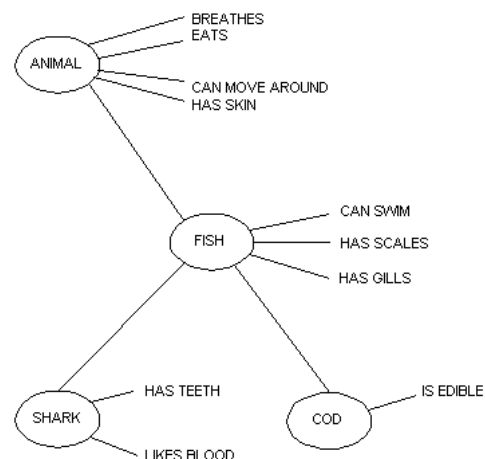
Básicamente, podemos distinguir tres categorías de redes semánticas:

1. Redes IS-A, en las que los enlaces entre nodos están etiquetados.
2. Grafos conceptuales: en los que existen dos tipos de nodos: de conceptos y de relaciones.
3. Redes de marcos: en los que los puntos de unión de los enlaces son parte de la etiqueta del nodo.

### 4.1. Redes IS-A

Se trata de la red semántica por excelencia. De hecho muchas veces se menciona este tipo como sinónimo de "red semántica", y los restantes tipos también incorporan el mismo tipo de enlaces o arcos (*links*).

Una red IS-A es una jerarquía taxonómica cuya espina dorsal está constituida por un sistema de enlaces de herencia entre los objetos o conceptos de representación, conocidos como *nodos*. Los otros tipos de redes semánticas son en realidad especializaciones de redes IS-A, por lo que siguen y amplían los conceptos fundamentales de las mismas.



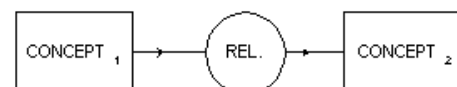
### 4.2. Grafos conceptuales

Los grafos conceptuales (*conceptual graphs*), se diferencian de las redes IS-A en que los arcos no están etiquetados, y los nodos son de dos tipos:

1. *Nodos de concepto*, que pueden representar tanto una entidad como un estado o proceso.
2. *Nodos de relación*, que indican cómo se relacionan los nodos de concepto.

Por tanto, los nodos de relación son los que hacen el papel de enlaces entre las entidades (los conceptos).

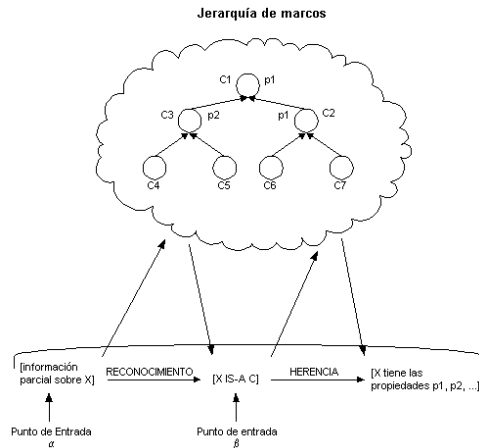
$$[\text{CONCEPT}_1] \Rightarrow (\text{REL.}) \Rightarrow [\text{CONCEPT}_2]$$



### 4.3. Esquemas de marcos

La fundamentación psicológica de los marcos es parecida a la de los *scripts*: cuando nos enfrentamos con una situación determinada, intentamos ajustarla a otra parecida de la que ya tenemos experiencia previa y *esperamos* que aparezcan un número de elementos comunes y se sucedan algunas situaciones.

Informalmente, un marco es una estructura de datos compleja que representa una situación estereotipada. Cada marco posee un número de casillas (*slots*) donde se almacena la información respecto a su uso y a lo que se espera que ocurra a continuación. Formalmente, un *marco*  $M$  es una lista de atributos (*slots*) y valores ( $S, V$ ). El valor  $V$  de un *slot*  $S$  puede ser un valor simple, cuyo tipo se halla determinado por el mismo *slot*, por una función que devuelve un valor de ese tipo, o por un puntero al valor de otro *slot* de la jerarquía; o una lista de atributos y valores, donde los atributos son funciones booleanas y los valores pueden ser funciones que devuelven valores.



## 5. BUSCADORES SEMÁNTICOS

Desde un punto de vista técnico, un buscador semántico es una aplicación que comprende las búsquedas de los usuarios y los textos de los documentos de la web mediante el uso de algoritmos que simulan comprensión o entendimiento, y que a partir de éstos proporciona resultados correctos sin que el usuario tenga que abrir el documento e inspeccionarlo por sí mismo. Un buscador de este tipo reconoce el contexto correcto para las palabras o sentencias de búsqueda.

Google y Yahoo! se basan fundamentalmente en algoritmos que generan estadísticas a partir de palabras y enlaces, no en algoritmos cognitivos que capturen el conocimiento implícito en las palabras y su contexto.

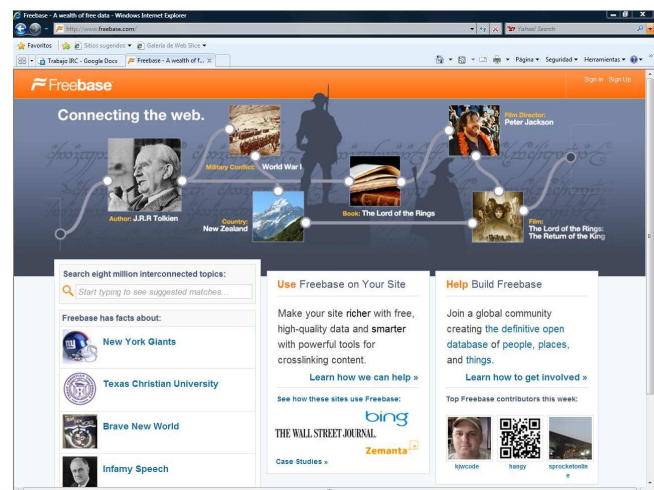
Por su parte, un buscador semántico no siempre puede acertar a la primera el significado que tiene una palabra dentro de una búsqueda, ya que las palabras pueden tener un distinto significado dentro de un contexto u otro, por lo que elegirán el significado más probable, pedirán directamente al usuario que elija entre varias opciones o utilizarán el resto de palabras que se incluyan en la búsqueda para ver el significado concreto de la palabra en su contexto.

El algoritmo de un buscador semántico se basa en simular la comprensión de las palabras y establecer relaciones entre ellas, y así, si por ejemplo alguien hace una búsqueda introduciendo las palabras "marsupial" el resultado serían documentos en los que apareciesen términos como canguro, koala, zarigüeya, demonio

de Tasmania,... De este modo uno puede encontrar documentos de interés que no se encontrarían buscando palabras clave.

La falta de estructura y de anotaciones semánticas en los recursos de la web (documentos Word, PDF, páginas HTML, etc.) obliga a que los buscadores semánticos analicen mediante algoritmos cognitivos los recursos, palabra a palabra y oración a oración, para asignar las palabras y oraciones a conceptos ontológicos. Estos algoritmos son lentos y requieren supervisión humana. De ahí que los buscadores semánticos no cubran por ahora tantos recursos de la web como los convencionales, que emplean algoritmos estadísticos, mucho más rápidos y completamente automatizados. Esta limitación desaparecerá cuando se vayan mejorando los algoritmos cognitivos o en cuanto los "islotos semánticos" se unan para formar la web semántica o, al menos, "continentes semánticos".

Existen ya buscadores que trabajan estructurando la información a la que luego se accede mediante búsquedas, ya que resulta posible obtener cierto nivel de conocimiento semántico a partir de la estructura de datos. Freebase, por ejemplo, es un buscador social que utiliza RDF (Resource Description Framework o Marco de Descripción de Recursos) para definir su estructura de datos como una serie de nodos y un conjunto de enlaces que establecen relaciones entre los nodos. Freebase tiene una gran colección de vocabularios codificados mediante RDF(S), con los que expresa un amplio rango de dominios y áreas de interés.



Lo que diferencia a Freebase de otras bases de datos es que cualquier asunto puede ir acompañado de muchas clases distintas de información. El ejemplo que dan es muy claro: "Por ejemplo, Arnold Schwarzenegger podría aparecer como actor en una base de datos de películas, como gobernador en una base de datos de política y como Mr. Universo en una base de datos de culturistas. En Freebase, solamente hay un tema sobre Arnold Schwarzenegger, que contiene información sobre las tres facetas de su vida pública. El tema unificado actúa como un centro de información, por lo que es fácil encontrar información sobre él y aportarla, independientemente de qué clase de información es.

En Freebase, cuando se accede a un objeto se tiene acceso a todos los aspectos de la información asociada



Miguel Ángel Abián, junio 2009

Por el momento, casi todos los buscadores semánticos permiten solamente búsquedas en inglés, aunque se están ampliando para que admitan otros lenguajes. Aparte del predominio del inglés, la causa de eso se debe también a las dificultades inherentes a reflejar el conocimiento de los lenguajes naturales en estructuras de datos que permitan búsquedas rápidas y escalables (matrices, listas, pilas, colas, árboles, grafos, etc.). Por ejemplo, el buscador Hakia utiliza un vocabulario en forma de ontología que incluye unos 100.000 sentidos de palabras inglesas, y ese número continuará aumentando según se perfeccione la aplicación.

Un buen ejemplo de cómo podrían ser los buscadores semánticos del futuro es Hakia, que permite búsquedas en lenguaje natural. Según la documentación oficial de Hakia, los resultados obtenidos en cualquier búsqueda satisfacen tres criterios:

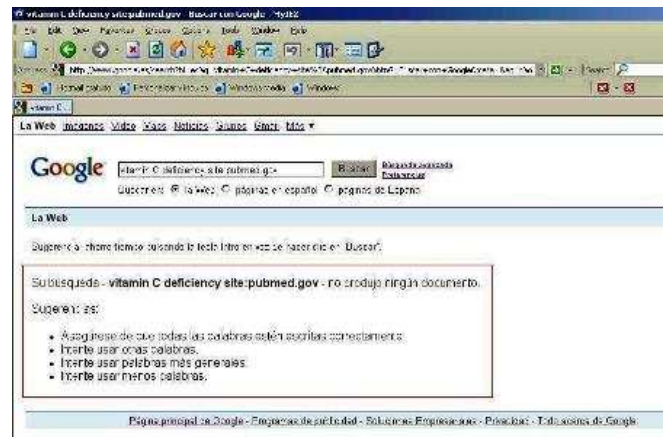
1. Proceden de sitios web creíbles (por ejemplo, en el caso de la salud, de sitios recomendados por la Medical Library Association).
2. Representan la información disponible más reciente.
3. Son absolutamente relevantes.

Los desarrolladores de Hakia rechazan el uso de estadísticas en un buscador verdaderamente semántico. Veamos el siguiente ejemplo: "Si tomamos la siguiente sentencia: *Los osos polares no comen huevos de caimanes antes del alba*. Estoy seguro de que usted nunca ha visto esta combinación de palabras antes. Pero el hecho de que pueda entender lo que significa es una prueba de que el cerebro no necesita muestreos estadísticos. El significado no emerge de la estadística. Emerge del conocimiento asociativo".

Google y Yahoo! no consideran mucha información relevante porque se centran únicamente en los contenidos más populares: ambos buscadores complementan sus algoritmos de búsqueda mediante palabras clave con estadísticas sobre el número de enlaces que apuntan a un recurso; este número, que mide la popularidad del recurso, determina su posición en la lista de resultados.

Como Hakia no cuenta enlaces para indexar los recursos, puede hacer visibles muchas páginas de interés con pocos enlaces o ninguno, que hasta ahora permanecen en la oscuridad o en la penumbra. Por caso, Hakia permite realizar búsquedas en unos 10

millones de artículos de PubMed.gov, un servicio de la Biblioteca Nacional de Medicina de los Estados Unidos que incluye resúmenes y citas de artículos de medicina. Cuando Google busca en ese servicio con los términos "vitamin D deficiency", devuelve 2 resultados, mientras que Hakia devuelve 49 (consulta realizada a mediados de julio), casi todos relevantes. En otros casos ("vitamin C deficiency", verbigracia), Google no devuelve resultados, mientras que Hakia sí. Esta disparidad se debe a las limitaciones inherentes a los métodos de búsqueda basados en estadísticas sobre los enlaces.



Las búsquedas semánticas que realiza Hakia emplean tres tecnologías. La primera es OntoSem, que procesa y analiza textos en lenguaje natural. Se basa fundamentalmente en una ontología, independiente del lenguaje, donde se almacenan miles de conceptos interrelacionados; y en un vocabulario o diccionario en inglés que almacena más de 100.000 sentidos de palabras, donde las palabras se categorizan según los distintos significados que tienen y donde se definen las relaciones entre ellas. Si se incluyen nombres propios, el diccionario cubre más de un millón de palabras inglesas. Este vocabulario enlaza los conceptos abstractos de la ontología con las palabras en inglés que se usan para denominarlos o referirse a ellos. La ontología cubre muchas las áreas de conocimiento, y es muy específica en áreas como medicina, ciencia y finanzas.

OntoSem incluye también un traductor o conversor ontológico (OntoParser) que traduce cada sentencia de un texto a una representación semántica (TMR o *Text Meaning Representation*) que usa los conceptos de la ontología y que se asemeja al significado que percibiría un hablante nativo. Por ejemplo, en una frase como "Saqué dinero del banco que hay en la esquina", el sentido de la palabra banco que proporcionaría el conversor sería "Establecimiento público de crédito, constituido en sociedad por acciones" y no "Asiento, con respaldo o sin él, en que pueden sentarse varias personas". OntoParser busca todas las acciones y sucesos con los posibles participantes de la frase que ha elegido antes. Luego da una puntuación a los posibles sucesos y acciones y a todas la combinaciones de sus participantes, teniendo en cuenta lo bien que éstos se acoplan a las acciones y sucesos.

La segunda tecnología en que se basa Hakia es QDEX (*Query indexing technique*). Esta técnica de indexación lee los



documentos (páginas HTML, por ejemplo) línea por línea y extrae todas las posibles búsquedas relativas al contenido de cada sentencia. Dicho de otro modo: QDEX anticipa todas las posibles consultas que pueden hacerse basándose en cada sentencia. Cuando un usuario escribe una pregunta en Hakia, esa pregunta figura ya en QDEX antes de que el usuario teclee nada; pues QDEX analizó las sentencias relevantes de los documentos de la Web y, basándose en tal análisis, dedujo anticipadamente que era una pregunta posible. (Los datos de QDEX se almacenan en formato semántico en una red distribuida de servidores, y el proceso de extracción de búsquedas es previo a cualquier consulta de los usuarios). Cuando alguien busca información, estas búsquedas —sentencias, a fin de cuentas— actúan como punteros a los documentos originales, los párrafos y las sentencias a partir de los cuales se generaron, de manera que se reducen los datos que deben manipularse al ejecutar una búsqueda concreta.

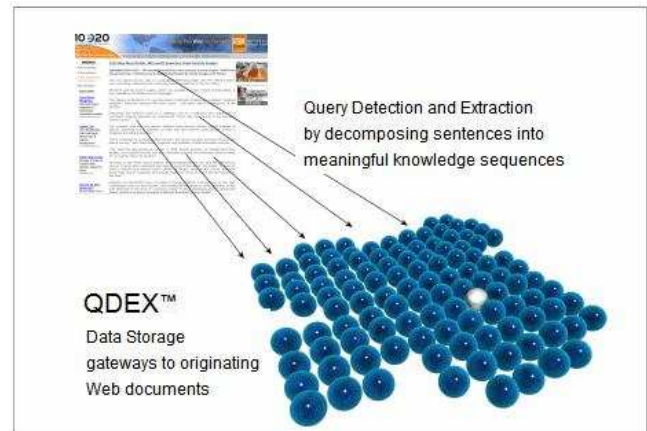
Cuando QDEX analiza los documentos, emplea OntoSem para eliminar las sentencias/búsquedas carentes de sentido.

Para profundizar en el funcionamiento de QDEX (*Query indexing technique*), hay que compararlo con los buscadores convencionales como Google, basados en índices invertidos de documentos, que permiten obtener una lista de documentos para una palabra clave. Un índice invertido es un índice que contiene palabras, que actúan como claves, emparejadas a referencias a los documentos donde aparecen. En los buscadores convencionales, los índices invertidos apuntan a una lista de identificadores de documentos y, dentro de cada identificador de documento, a una o más posiciones dentro de él.

Las búsquedas basadas en índices invertidos exigen mantener activa una gran cantidad de conjuntos de datos antes de cualquier consulta. En las búsquedas semánticas, esto constituye un grave problema: hay que *enriquecer semánticamente* los datos con relaciones entre conceptos, lo cual ocasiona que se multiplique exponencialmente la memoria que debe destinarse a almacenarlos y el tiempo de cada búsqueda.

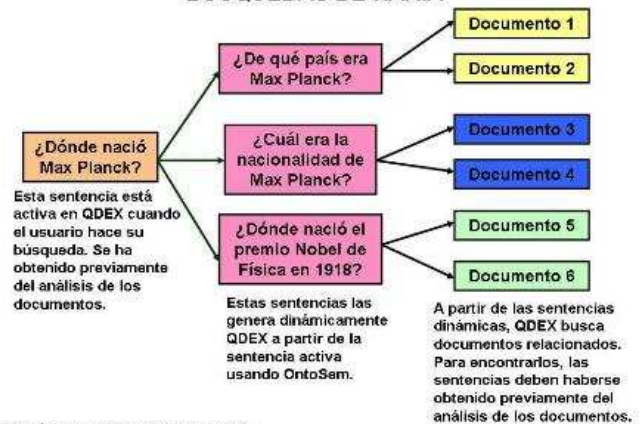
Por caso, un índice invertido como "cerrar" debería, al enriquecerse semánticamente con conceptos como aislar, abrochar, sellar, estrechar, obstruir, bloquear..., mantener en memoria la relación con los índices invertidos correspondientes a esos conceptos. Cuando se buscara con la palabra *acabar*, el buscador semántico recorrería las relaciones y devolvería tanto los documentos con la palabra anterior como aquellos con *aislar*, *abrochar*, etc.

En lugar de tener una larga serie de índices invertidos activos antes de cualquier búsqueda, QDEX mantiene activo un pequeño conjunto de sentencias o búsquedas ya preparadas para cada posible pregunta y obtiene dinámicamente el resto mediante asociaciones semánticas a partir de OntoSem.



Cuando el usuario formula alguna pregunta, QDEX recurre a OntoSem para generar sentencias equivalentes a la pregunta o relacionadas con ella. Usando estas sentencias, devuelve luego los documentos de interés. Lógicamente, si las sentencias generadas no están almacenadas previamente en QDEX —y, por tanto, vinculadas a documentos—, no se devuelve ningún documento. Veamos el funcionamiento con un ejemplo gráfico:

#### EJEMPLO DE FUNCIONAMIENTO DE QDEX EN LAS BÚSQUEDAS DE HAKIA

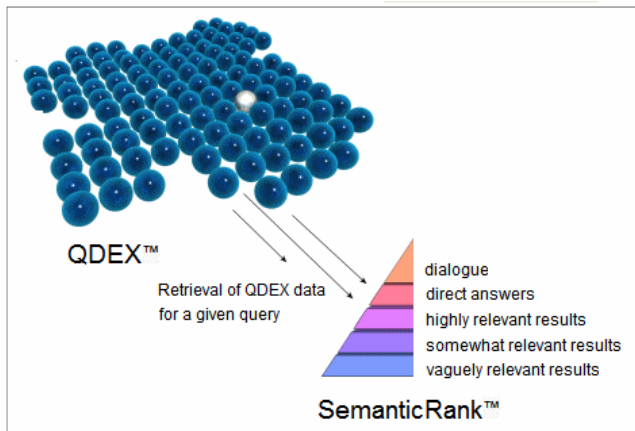


Miguel Ángel Abián, septiembre 2009

Mantener activo un reducido número de búsquedas permite que Hakia responda rápidamente a las consultas. QDEX genera dinámicamente sentencias a partir de las activas, encuentra documentos relevantes aunque no contengan las palabras exactas de la búsqueda.

Por último, Hakia se basa en el algoritmo SemanticRank, que toma como entrada los párrafos relevantes que proceden de QDEX para una determinada búsqueda. Este algoritmo determina la relevancia de los resultados que se van a mostrar, y por tanto su orden, basándose en un análisis de las concordancias entre los conceptos relacionados con la búsqueda (no solamente las palabras clave) y las sentencias más relevantes de cada párrafo. Además, se tiene en cuenta la credibilidad y la antigüedad de los contenidos para determinar su relevancia (por ejemplo, una

página web mal escrita o estructurada tendrá poca o ninguna relevancia). A diferencia de Google, SemanticRank no tiene en cuenta el número de enlaces entre documentos, sino que analiza el contenido.

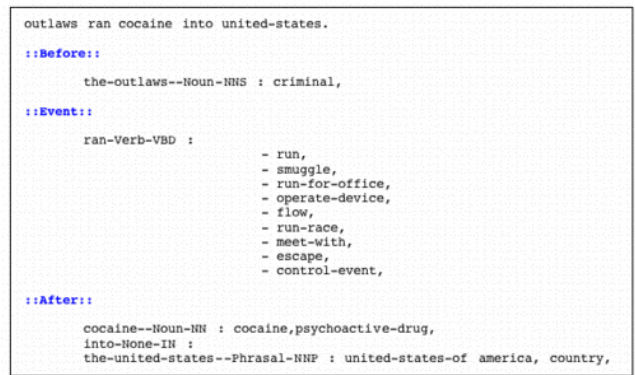


## 6. EJEMPLO DE FUNCIONAMIENTO DE HAKIA

En la propia página de Hakia podemos encontrar el siguiente ejemplo, de lo hace Hakia para realizar una búsqueda. La frase que utiliza es la siguiente: "Outlaws ran cocaine into United States" ("Bandidos pasan de contrabando cocaína en los Estados Unidos").

Como hemos explicado antes, Hakia pasa por tres fases OntoSem, Qdex y SemanticRank.

OntoSem incluye un conversor ontológico (OntoParser) que traduce cada sentencia de un texto a una representación semántica (TMR o *Text Meaning Representation*). OntoParser genera todos los potenciales sentidos de las palabras de la oración y la rompe en partes que se basan en las acciones o sucesos principales que se identifican a partir de esos sentidos.

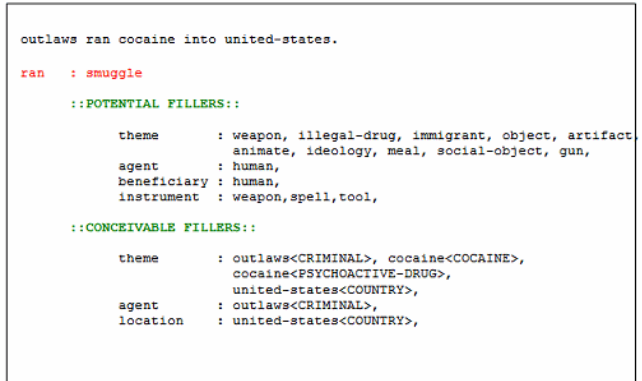


La palabra "outlaw" (bandido, forajido) tiene un solo sentido en inglés: CRIMINAL; "ran", pretérito del verbo "run" (correr, tomar parte en, funcionar, ejecutar, dirigir, organizar, transportar, pasar

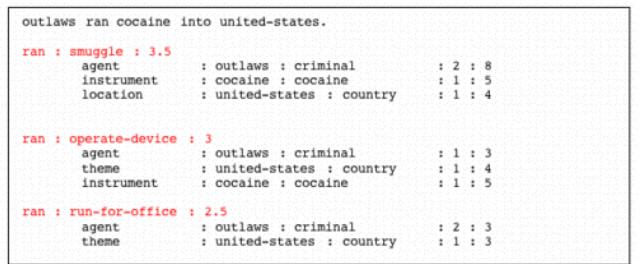
de contrabando [armas o drogas]), tiene 9 sentidos en inglés; "cocaine", 2; y "United States", uno.

No todos los sentidos de "ran" son válidos. Por ejemplo, FLOW (fluir) no admite agentes, solamente un tema (no se puede decir, p. ej., "Andrea fluyó el mercurio"), y debe ser un líquido; y ni CRIMINAL ni DRUG (droga, medicamento) son líquidos.

La acción SMUGGLE (contrabandear, pasar de contrabando) solamente admite temas que deben ser WEAPON (arma), ILLEGAL-DRUG (droga ilegal) o IMMIGRANT (inmigrante), como muestra la siguiente imagen.



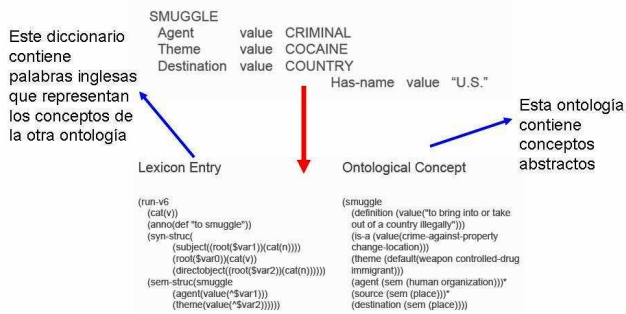
OntoParser rellena todas las acciones y sucesos con los posibles participantes de la frase que ha elegido antes. Luego da una puntuación a los posibles sucesos y acciones y a todas las combinaciones de sus participantes, teniendo en cuenta lo bien que éstos se acoplan a las acciones y sucesos. SMUGGLE es la opción más correcta (3,5) para el significado de "ran" porque COCAINE, una ILLEGAL-DRUG, es lo que mejor encaja con el tema que SMUGGLE puede tomar. Las otras dos opciones se revelan erróneas para cualquier ser humano cabal, pero no para una máquina: ¿cómo van los bandidos a hacer funcionar los Estados Unidos con cocaína o cómo van a ser candidatos (*run for office*) a Estados Unidos?



La representación final en TMR (Text Meaning Representation) o representación del significado del texto, es lo que finalmente se usará para realizar la búsqueda.

## Outlaws ran cocaine into the country

Representación semántica de la frase de arriba  
TMR = Text meaning representation



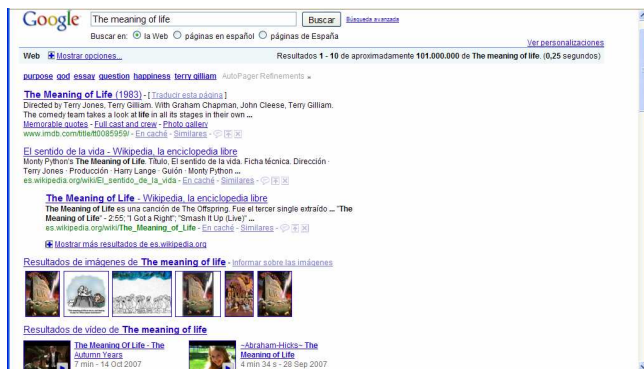
Ejemplo tomado de Cyber Security and Information Intelligence Research Workshop, Mayo 2008  
Oak Ridge National Laboratory

Si existiera la Web semántica, Hakia no necesitaría OntoParser: buscaría directamente en las representaciones semánticas de las páginas electrónicas, de manera similar a lo que hace Swoogle con los documentos RDF que encuentra en la Red. Como no es así, primero debe transformarse el lenguaje natural de los recursos en una representación formal.

## 7. COMPARACIÓN ENTRE GOOGLE Y HAKIA

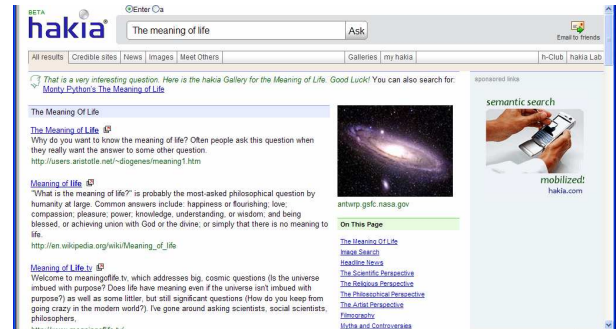
Como ya se ha explicado en este documento, un buscador semántico tiene una mayor precisión en cuanto resultado de las búsquedas que realiza el usuario. Aquí podemos ver un ejemplo entre Google y Hakia realizando la misma búsqueda en ambos. La búsqueda es: *"The meaning of life"*.

Mientras que Google es un buscador de palabras clave y nos da como resultado las páginas más visitadas en las cuales se encuentran alguna o varias de las palabras usadas en la búsqueda, Hakia nos da como resultado, páginas en las que se intenta dar una explicación del "Significado de la vida".



En Google lo primero que encontramos son páginas referidas a la película de Monty Python que tiene ese mismo título: The meaning of life. Más adelante nos sugiere imágenes también de

dicha película y unos videos. Despues, a pesar de no verse en la imagen, nos encontramos con páginas de tipo más filosófico o enlaces a páginas de enciclopedias libres como wikipedia, en las que se ha encontrado la frase literalmente.



En Hakia lo que vemos son páginas donde puede que esté literalmente o no la misma frase, pero el contenido tiene que ver precisamente con la búsqueda. Encontramos artículos y textos filosóficos acerca de la vida y su significado. También más adelante (aunque en la imagen no se puede apreciar), nos da todas las posibles facetas de la búsqueda que nos podrían interesar (The Scientific Perspective, The Religious Perspective, The Philosophical Perspective, The Artist Perspective, Filmography).

Como hemos podido comprobar, en los buscadores semánticos tenemos la facilidad de realizar búsquedas en lenguaje natural, no como nos ha pasado a todos muchas veces de pensar: "¿Qué otra palabra puedo añadir a la búsqueda para cerrar más el círculo y no tener que abrir 100 páginas para dar solo con una o dos que merezcan la pena?"

## 8. REFERENCIAS

[1] Miguel Ángel Avián (JavaHispano), 1993. [en línea] [ref. de 20 de Septiembre de 2005]  
<<http://www.javahispano.org/tutorials.item.action?id=55>>

[2] Miguel Ángel Avián (Web semántica hoy), 2009. [en línea] [ref. de 10 de Julio de 2009]  
<<http://www.wshoy.sidar.org/index.php?2009/07/10/47-buscadores-semanticos-hakia-por-dentro-y-por-fuera-parte-2>>

[3] Autor desconocido(Ciencia.net), 2004. [en línea] [ref. de 9 de Abril de 2004 ]  
<<http://www.ciencia.net/VerArticulo/?idTitulo=Ontolog%EDA%20%28Inform%Etica%29>>

[4] Autor desconocido, Fecha desconocida. [en línea]  
<<http://iteso.mx/~carlosc/pagina/cursoUAHI/Conferencias/buscadores.htm>>

[5] Wikipedia (Enciclopedia de contenido libre). Motor de búsqueda. [en línea] [ref. de 1 de Diciembre de 2009]  
<[http://es.wikipedia.org/wiki/Motor\\_de\\_b%C3%BAsqueda](http://es.wikipedia.org/wiki/Motor_de_b%C3%BAsqueda)>

[6] Wikipedia (Enciclopedia de contenido libre). Redes semánticas. [en línea] [ref. de 2 de Octubre de 2009]

<[http://es.wikipedia.org/wiki/Redes\\_sem%C3%A1nticas](http://es.wikipedia.org/wiki/Redes_sem%C3%A1nticas)>

[7] Blog de Hakia.com [en línea] [ref. de 17 de Febrero de 2008 ]

<<http://blog.hakia.com/?p=228>>

[8] Swoogle (Web de búsqueda semántica), 2007

<<http://swoogle.umbc.edu>>

[9] Antonio Moreno Ortiz, Diseño e implementación de un lexicón computacional para lexicografía y traducción automática. Representación del conocimiento y representación léxica, 2000 [en línea]

<<http://elies.rediris.es/elies9/4-3-2.htm>>